Journal of the American Statistical Association

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/uasa20

Efficient Estimation of Optimal Regimes Under a No Direct Effect Assumption

Lin Liu , Zach Shahn , James M. Robins & Andrea Rotnitzky

To cite this article: Lin Liu , Zach Shahn , James M. Robins & Andrea Rotnitzky (2021): Efficient Estimation of Optimal Regimes Under a No Direct Effect Assumption, Journal of the American Statistical Association, DOI: 10.1080/01621459.2020.1856117

To link to this article: https://doi.org/10.1080/01621459.2020.1856117

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION
2021, VOL. 00, NO. 0, 1–16: Precision Medicine
https://doi.org/10.1080/01621459.2020.1856117

Efficient Estimation of Optimal Regimes Under a No Direct Effect Assumption

Lin Liu[a], Zach Shahn[b], James M. Robins[c], and Andrea Rotnitzky[d]

[a]Institute of Natural Sciences, School of Mathematical Sciences, MOE-LSC and SJTU-Yale Center for Biostatistics and Data Science, Shanghai Jiao Tong University, Shanghai, China; [b]IBM Research, Yorktown Heights, NY; [c]Department of Biostatistics and Epidemiology, Harvard University, Boston, MA; [d]Department of Economics, Universidad Torcuato Di Tella and CONICET, Buenos Aires, Argentina

ABSTRACT

We derive new estimators of an optimal joint testing and treatment regime under the no direct effect (NDE) assumption that a given laboratory, diagnostic, or screening test has no effect on a patient's clinical outcomes except through the effect of the test results on the choice of treatment. We model the optimal joint strategy with an optimal structural nested mean model (opt-SNMM). The proposed estimators are more efficient than previous estimators of the parameters of an opt-SNMM because they efficiently leverage the "NDE of testing" assumption. Our methods will be of importance to decision scientists who either perform cost-benefit analyses or are tasked with the estimation of the "value of information" supplied by an expensive diagnostic test (such as an MRI to screen for lung cancer). Supplementary materials for this article are available online.

ARTICLE HISTORY
Received July 2019
Accepted September 2020

KEYWORDS
Direct effects; Nuisance tangent space; Optimal dynamic regimes; Optimal regime structural nested mean models; Screening

1. Introduction

This article provides new, highly efficient estimators of an optimal joint testing and treatment regime under the no direct effect (NDE) assumption that a given laboratory, diagnostic, or screening test has no effect on a patient's clinical outcome of interest except through the effect of the test results on the choice of treatment. The proposed estimators attain high efficiency because they leverage this "NDE of testing" assumption. What is surprising and, indeed, unprecedented in the experience of the authors, is that, in actual substantive studies, estimators that leverage the NDE assumption can provide a 50-fold increase in efficiency (and, thus, a 50-fold reduction in sample size) compared to estimators that fail to leverage the NDE assumption (Caniglia et al. 2019).

The following case study of a randomized trial in HIV infected subjects in Africa motivates the issues with which we are concerned. First some background. HIV infected subjects who start on first line antiretroviral therapy (ART) and who later develop evidence of disease progression, quantified by an increase in viral load, a decrease in CD4 immune cell count in blood, or new clinical symptoms must consider switching to second line therapy. Therefore, at each clinic visit, a decision based on the previous laboratory and clinical data must be made as to (1) whether to order viral load and/or CD4 count tests at some cost and burden to the patient and (2) whether to switch treatment. Waiting too long to switch may result in further clinical deterioration and eventually AIDS and/or death. On the other hand, switching too early is unwise as there are only a limited number of alternative regimes and these are associated with increased side effects and expense. Thus, there is a need to determine the optimal testing and treatment regime based on empirical analysis of observational or randomized trial data. Tests of viral load and CD4 count in themselves have no direct biological effect on a patient. Rather the test results are used to help gauge the likely benefit of switching treatment. Hence appropriately timed tests may result in an increase in the expected utility, even if the financial and other costs of the test are considered.

The issue of how to determine when to switch to second line therapy is most pressing in Africa, where there has been a longstanding concern that first line ART is prescribed with limited, often no, laboratory monitoring of disease progression. Therefore, in 2003, HIV infected individuals were recruited into the Development of AntiRetroviral Therapy in Africa (DART) randomized trial in which individuals on first line ART in Uganda and Zimbabwe were randomized to laboratory and clinical monitoring (LCM) versus clinically driven monitoring (CDM). In the LCM arm, CD4 cell counts were measured every 12 weeks and made available to the caregivers. In the CDM arm, CD4 counts were not available to the caregivers. LCM subjects were switched to second line therapy when either the CD4 count fell below 100 or a world health organization (WHO) stage 4 clinical event occurred; CDM arm subjects were switched at the occurrence of a WHO stage 4 event. The trial results were published by the DART Team in 2010 (DART Trial Team 2010). An intention to treat analysis showed an adverse effect of CDM compared to LCM both for overall mortality (hazard ratio 1.35, 95% CI [1.10, 1.65]) and for the primary outcome, time to (the minimum of) a WHO stage 4 event and death (hazard ratio 1.31, 95% CI [1.14, 1.51]).

CONTACT James M. Robins *robins@hsph.harvard.edu* Department of Biostatistics and Epidemiology, Harvard University, Boston, MA 02115.
Supplementary materials for this article are available online. Please go to *www.tandfonline.com/r/JASA*.

In 2012, the DART Team published a cost-effectiveness analysis (Lara et al. 2012). Owing to both the cost of CD4 count testing every 12th week and the higher price of second line therapy, their analysis estimated an incremental cost of at least 3000 US dollars for each additional quality-adjusted life year (QALY) gained from following the LCM strategy compared to the CDM strategy. By its definition, a cost effectiveness analysis compares the ratio of incremental dollar costs to the health benefit in QALYs for a number of different health interventions. The philosophy underlying this definition is that those interventions, ranked from the lowest to highest ratio, would be implemented in that order until available funds are exhausted. (This is a simplification, overlooking, for instance, that factors (individual and social) other than QALYs may be used to quantify the benefits of an intervention.) An alternative approach is to combine the costs and benefits into a single utility function whose expected value we wish to maximize. An analysis whose goal is to maximize a single utility function is generally referred to as a cost-benefit, rather than a cost-effectiveness, analysis. A cost-benefit analysis requires that we place a monetary value on an additional year of quality-adjusted life, as discussed next.

The DART Team noted that a WHO publication (World Health Organization 2001) suggested that, to be cost effective, the incremental cost of a proposed intervention per QALY added should be less than three times the per capita GDP. If one takes the WHO literally (which may not be what WHO meant for one to do), it allows one to place an upper bound on the monetary value of a year of quality-adjusted human life. For example, the per capita GDP of Zimbabwe and Uganda were both less than 480 US dollars in 2008 (presumably the last year available to the authors). The DART Team thus concluded that the LCM strategy was not cost effective compared to the CDM strategy as $3 \times 480 = 1440$ is less than incremental cost of 3000 dollars. That is, they found that the monetized health value of the information (VoI) concerning disease progression obtained from CD4 testing every 12th week was less than the cost of obtaining the information. Thus, in terms of the definitions above, the DART Team had performed a cost-benefit analysis rather than simply a cost-effectiveness analysis. In contrast, the 2008 GDP/per capita of the United States was 48,000 dollars. Hence, the WHO criteria, taken literally, would suggest that the monetary value of a QALY for a US citizen was 144,000 dollars, 100 times that of a citizen of Uganda or Zimbabwe. We leave further discussion of the obvious ethical and economic issues raised by monetizing a human life to others. In this article, we will assume the cost of testing and the health benefit are combined into a single utility function. A natural question that arises is whether the above WHO cost-effectiveness criterion would be met if, instead of testing every 12 weeks as in the trial, testing occurred every 24 weeks or 36 weeks or 48 weeks with switching to second line therapy when the CD4 count first fell below 50, 100, 150, or 200. None of these $3 \times 4 = 12$ testing and treatment regimes were studied in the RCT so any attempt to answer such questions requires one analyze the trial data as if it were an observational study using methods related to those proposed herein (Ford et al. 2015).

Finding an optimal joint testing and treatment regime for HIV infected subjects is but one of many contexts in which our methodology should be applicable. Many laboratory, diagnostic and screening tests have no effect on the clinical outcome of interest, except through the effect of the test results on the choice of treatment. An area in which our new, more efficient estimators should be particularly important is that of cost-benefit analyses wherein the costs of expensive tests (such as MRIs to screen for lung cancer, mammograms to screen for breast cancer, and urinary cytology to screen for bladder cancer) are weighed against the clinical VoI supplied by the test results (e.g., Mushlin and Fintor 1992; Krahn et al. 1994).

There is a large literature in economics and decision science on the VoI. For the purposes of this article, we define the VoI to be the increase in the expected utility resulting from incorporating costly information without a direct causal effect on the outcomes of interest (such as a screening test) into an optimal regime (e.g., LaValle 1968a, 1968b; Gould 1974; Hilton 1981). That is, the VoI is the difference in expected utilities of two regimes: the optimal testing and treatment regime versus the optimal treatment regime under the constraint that no testing is allowed. Our methodology also allows for more efficient estimation of the VoI.

In fact, the VoI can be directly calculated from the parameters of an optimal regime structural nested mean models (opt-SNMM). Robins (2004), building on Murphy (2003), introduced the opt-SNMM, a semiparametric model for estimating the optimal testing and treatment regime from data. Under the model, the optimal testing and treatment regime is a deterministic function of the model parameters Ψ.

Robins (2004) proposed g-estimation, a semiparametric version of dynamic programming (DP), to estimate the parameters Ψ^* of an opt-SNMM. Under standard assumptions required for the identification of causal effects in longitudinal settings (consistency, positivity, and sequential exchangeability), g-estimation of a correctly specified opt-SNMM yields a regular, asymptotically linear (RAL) estimator $\widetilde{\Psi}$ of Ψ^* and thus of the optimal joint testing and treatment regime and its value (i.e., expected utility), provided the true law generating the data is not an exceptional law as defined in Robins (2004, p. 219) and the bias of $\widetilde{\Psi}$ is of order $o_p(n^{-1/2})$ with n the sample size. In this article, we exclude such exceptional laws for reasons given later in this section. We discuss conditions required for the bias to be $o_p(n^{-1/2})$ in Appendix A.5. An estimator $\widetilde{\Psi}$ is RAL if (i) $\widetilde{\Psi} - \Psi^*$ is the sum of iid mean zero random variables (referred to as the influence function of $\widetilde{\Psi}$) plus a term of order $o_p(n^{-1/2})$ and (ii) is locally asymptotically unbiased.

In this article, we show that under the NDE assumption it is possible to construct RAL estimators that are more efficient than the g-estimators of Robins (2004) that do not use the NDE assumption. Our construction is not straightforward because imposing the NDE assumption does not restrict the values of any of the parameters of our opt-SNMM, including those parameters that determine the optimal testing regime. For a more comprehensive review of SNMMs, we refer interested readers to the articles by Robins (2000, 2004) or a more recent piece by Vansteelandt and Joffe (2014).

We now briefly describe our estimator construction. Full details are given in Sections 3–6. Robins (2004) showed that, without imposing the NDE assumption, the g-estimator $\widetilde{\Psi}$ was equal to the solution of estimating equations $0 = \widehat{\mathbb{U}}(\Psi)$ where,

at the true Ψ^* , $\widetilde{\Psi}$, and $\widehat{\mathbb{U}}(\Psi^*)$ were RAL estimators of Ψ^* and 0, respectively, and, in addition, were doubly robust in the sense of Bang and Robins (2005); see Section 3 for further discussion. The new estimators $\widetilde{\Psi}(b)$ in this article solve $0 = \widehat{\mathbb{U}}(\Psi, b)$ where $\widehat{\mathbb{U}}(\Psi, b)$ is the residual from the orthogonal projection of the (influence function of) $\widehat{\mathbb{U}}(\Psi)$ into a given linear space of random variables T_b indexed by a vector function b. In the absence of confounding by unmeasured factors, all components of T_b have mean zero under the NDE assumption. To construct T_b, we first show that the NDE assumption implies the following restriction on the distribution of the observed data. Given past history, the screening variable at any time t is independent of the health outcome of interest after reweighting by the inverse of the conditional probability of a subject's treatment history from $t + 1$ onward. We then use the independence of screening and the health outcome in this reweighted distribution to construct the vector T_b. The Pythagorean theorem then guarantees that the asymptotic relative efficiency (ARE) $\lim_{n\to\infty} \text{var}(\widetilde{\Psi})/\text{var}(\widetilde{\Psi}(b))$ of the old relative to the new estimator is always greater than or equal to 1. Furthermore the new estimator, like the old, is doubly robust.

In the article, we assume a semiparametric model on the joint distribution of the factual and counterfactual variables defined by the restrictions that (i) the NDE assumption is true, (ii) confounding by unmeasured variables is absent, and (iii) the opt-SNMM holds. The regular estimator $\widetilde{\Psi}(b_{\text{opt}})$ with minimum asymptotic variance (and hence the notation b_{opt}) among all RAL estimators under the model solves $0 = \widehat{\mathbb{U}}(\Psi, b_{\text{opt}})$ where $\widehat{\mathbb{U}}(\Psi, b_{\text{opt}})$ is the residual from the projection $T_{b_{\text{opt}}}$ (indexed by b_{opt}) of (the influence function of) $\widehat{\mathbb{U}}(\Psi)$ onto the space of all random variables with mean zero under and only under the NDE assumption. We show in Section 5.2 that, when the utility is a continuous random variable, this projection and thus $\widetilde{\Psi}(b_{\text{opt}})$ depend on the observed data distribution through the solutions to a set of complicated integral equations that do not exist in closed form. In contrast, when the utility is discrete, we obtain a closed form expression for the projection. Furthermore, in the case of a continuous utility (see Remark 4), we propose a *computationally tractable* estimator $\widetilde{\Psi}(b_{\text{sub}})$ solving $0 = \widehat{\mathbb{U}}(\Psi, b_{\text{sub}})$ with relative efficiency that can be made arbitrarily close to that of the "optimal" yet *computationally intractable* estimator $\widetilde{\Psi}(b_{\text{opt}})$. Specifically $\widehat{\mathbb{U}}(\Psi, b_{\text{sub}})$ is the residual from the closed-form projection $T_{b_{sub}}$ (indexed by b_{sub}) of (the influence function of) $\widehat{\mathbb{U}}(\Psi)$ onto a large, but strict, subspace of the space of random variables with mean zero under the NDE assumption. As the dimension of the chosen subspace increases, the relative efficiency of $\widetilde{\Psi}(b_{\text{sub}})$ approaches that of $\widetilde{\Psi}(b_{\text{opt}})$; hence if we allow the dimension to converge to infinity with the sample size at a sufficiently slow rate, $\widetilde{\Psi}(b_{\text{sub}})$ and $\widetilde{\Psi}(b_{\text{opt}})$ will be asymptotically equivalent. Because the projection needed to compute $\widetilde{\Psi}(b_{\text{sub}})$ exists in closed-form, it is much easier to compute than $\widetilde{\Psi}(b_{\text{opt}})$ and is therefore the estimator we recommend.

We shall study the relative efficiency of the estimators $\widetilde{\Psi}(b_{\text{sub}})$ and $\widetilde{\Psi}$ that, respectively, do and do not use the NDE assumption, as estimators of the parameters Ψ^* of an opt-SNMM. However, our approach is quite generic in the following sense. Consider a semiparametric model for the joint distribution of the factual and counterfactual variables defined by the restrictions that (i) the NDE assumption is true, (ii) confounding by unmeasured variables is absent, and (iii) a given semiparametric model holds with Ψ^* the true value of the Euclidean parameter. As one important example other than an opt-SNMM, the given model could be a dynamic marginal structural model (dyn-MSM) for a prespecified class of testing and treatment regimes (Hernán 2005; van der Laan and Petersen 2007; Robins, Orellana, and Rotnitzky 2008; Orellana, Rotnitzky, and Robins 2010a, 2010b). Given a doubly robust RAL estimator $\widetilde{\Psi}$ of Ψ^* that solves $0 = \widehat{\mathbb{U}}(\Psi)$ (with $\widehat{\mathbb{U}}(\Psi)$ an unbiased estimating equation in the semiparametric model in the absence of the NDE assumption), the methods in this article can be used to construct a RAL doubly robust estimator $\widetilde{\Psi}(b_{\text{sub}})$ with improved efficiency compared to $\widetilde{\Psi}$ by solving $0 = \widehat{\mathbb{U}}(\Psi, b_{\text{sub}})$ where again $\widehat{\mathbb{U}}(\Psi, b_{\text{sub}})$ is the residual from the closed-form projection (indexed by b_{sub}) of (the influence function of) $\widehat{\mathbb{U}}(\Psi)$ onto the exact same subspace as earlier. The only difference is that $\widehat{\mathbb{U}}(\Psi)$ and its influence function differ depending on the chosen semiparametric model (e.g., opt-SNMM vs. dyn-MSM) for Ψ^*. Since our procedure is a generic procedure applied to an initial RAL estimator $\widetilde{\Psi}$ and we are simply using an opt-SNMM as a particular example, we decided, as mentioned above, to exclude exceptional laws because the opt-SNMM g-estimator $\widetilde{\Psi}$ is not a RAL estimator under an exceptional law.

The methodology discussed above constructs highly efficient estimators under the NDE assumption of NDE of testing on the clinical outcome Y^d of interest. However even more efficient estimators can be constructed if we can impose the stronger NDE assumption of NDE of testing not only on Y^d but also on measured time dependent covariates, as the stronger NDE assumption implies additional mean zero random variables on which to project. See Section 7 for more discussion.

The above development offers the reader little guidance or intuition as to when to expect small versus enormous gains in efficiency compared to estimators that do not rely on the NDE assumption. To provide intuition, in Section 8, we discuss an inverse probability weighting (IPW) estimator $\widetilde{\Psi}_{\text{nde-ipw}}$ introduced in Robins, Orellana, and Rotnitzky (2008) and used by Caniglia et al. (2019) in a substantive article that uses a dyn-MSM to determine the optimal testing and treatment strategy [within a class of (CD4 cell and HIV RNA) testing and (antiretroviral) treatment strategies] to prolong the survival of HIV infected individuals included in the Harvard HIV-CAUSAL Collaboration and the Centers for AIDS Research Network of Integrated Clinical Systems. The estimator $\widetilde{\Psi}_{\text{nde-ipw}}$ has also been studied by Neugebauer et al. (2017) and Kreif et al. (2020). $\widetilde{\Psi}_{\text{nde-ipw}}$ is an asymptotically linear, though generally inefficient, estimator under the stronger NDE assumption; however, it is inconsistent when this assumption is false. In the analysis of Caniglia et al. (2019), $\widetilde{\Psi}_{\text{nde-ipw}}$ is unprecedentedly 50 times as efficient as the usual IPW estimator $\widetilde{\Psi}_{\text{ipw}}$ that does not exploit an NDE assumption. In fact, $\widetilde{\Psi}_{\text{nde-ipw}}$ would be nearly 50 times as efficient as any estimator that remains RAL when the NDE assumption does not hold.

Even so, we show in Section 8.1 that $\widetilde{\Psi}_{\text{nde -ipw}}$ will still be less efficient than $\widetilde{\Psi}_{\text{ipw}}(b_{\text{opt, ipw}})$. Here $\widetilde{\Psi}_{\text{ipw}}(b_{\text{opt, ipw}})$ solves $0 = \widehat{\mathbb{U}}_{\text{IPW}}(\Psi, b_{\text{opt, ipw}})$ where $\widehat{\mathbb{U}}_{\text{IPW}}(\Psi, b_{\text{opt, ipw}})$ is the residual from the projection (indexed by $b_{\text{opt, ipw}}$) of the estimated influence function of $\widetilde{\Psi}_{\text{ipw}}$ onto the space of all random variables with mean zero only under the stronger NDE assumption! Indeed, we show that $\widetilde{\Psi}_{\text{ipw}}(b_{\text{opt, ipw}})$ is semiparametric efficient in the model characterized solely by the stronger NDE assumption if the dyn-MSM is a saturated model.

In Section 9, we design a simple data generating process (DGP) that makes it transparent why $\widetilde{\Psi}_{\text{nde-ipw}}$ and $\widetilde{\Psi}_{\text{ipw}}(b_{\text{opt, ipw}})$ are so much more efficient than $\widetilde{\Psi}_{\text{ipw}}$ in the context of Caniglia et al. (2019). The above may lead one to suspect that $\widetilde{\Psi}_{\text{nde-ipw}}$ is always more efficient than the usual IPW estimator $\widetilde{\Psi}_{\text{ipw}}$, whenever the NDE assumption is imposed. However this is far from the case. In fact, we design a second DGP (see Appendix A.12) under which, as a particular parameter of the DGP converges to 0, the ratio of the asymptotic variance of $\widetilde{\Psi}_{\text{ipw}}$ compared to $\widetilde{\Psi}_{\text{nde-ipw}}$ tends to 0 and simultaneously the asymptotic variance of $\widetilde{\Psi}_{\text{ipw}}$ converges to that of the efficient estimator $\widetilde{\Psi}_{\text{ipw}}(b_{\text{opt, ipw}})$.

The organization of the article is as follows. In Section 2, we establish some notation and review the counterfactual framework. In Section 3, we review opt-SNMMs and g-estimation without imposing the NDE assumption. In Section 4, we characterize the ortho-complement of the tangent spaces under the NDE assumption as a key step toward constructing more efficient estimators. In Section 5, we describe several strategies for obtaining more efficient estimators by leveraging the NDE assumption. In particular, we describe the aforementioned computationally tractable procedure used in construction of $\widetilde{\Psi}(b_{\text{sub}})$ in Section 5.2. In Section 6, we study the statistical properties of the proposed estimators when the nuisance parameters/functions are estimated from data possibly using machine learning methodology. In Section 7, we introduce increasingly strong versions of the NDE assumption and compare their substantive plausibility. In Section 8, we provide the intuition behind our large efficiency gains through a study of the properties of $\widetilde{\Psi}_{\text{nde-ipw}}$. We also formally compare the statistical properties of all the various estimators treated in the article. In Section 9 and Appendix A.9, we illustrate the method on simulated examples. In Section 10, we conclude with some open problems and future directions.

2. Notation, Framework, and Background

We begin by providing the notation that will be used throughout the article. Let:

- $t \in \{0, \ldots, K\}$ index time or visits, assumed discrete, with K being the last occasion;
- $A_t \in \{0, 1\}$ denote whether a screening test is performed at time t;
- R_t denote the results of the test (e.g., MRI screening result) if performed at $t-1$ and $R_t =?$ otherwise;
- S_t be a binary $\{0, 1\}$ variable denoting the treatment (e.g., switching therapy) at time t;
- L_t denote covariates at time t that may influence testing and treatment decisions;
- Y^d denote the observed value of the health outcome utility;
- $Y \equiv Y^d - \sum_{t=0}^{K} c^* A_t$ denote the total utility of interest with c^* being the known (utility) cost of the test at time t; and
- $\bar{X}_t$ denote $(X_0, \ldots, X_t)$ and $\underline{X}_t$ denote $(X_t, \ldots, X_K)$ for an arbitrary vector $X \equiv (X_0, \ldots, X_K)$.

We assume that we observe N iid realizations of the random vector:

$$O \equiv (L_0, R_0, S_0, A_0, \ldots, L_K, R_K, S_K, A_K, Y^d).$$

We use capital letters to denote random variables and corresponding lower case letters to denote specific values that random variables might take. We consider the scenario where at each time t the chronological ordering of the variables is L_t before R_t before S_t before A_t. Our definition of R_t implies that the results of tests at time t are not available until time $t+1$. (By convention, we take A_K to always be 0 as results of testing at K would not be available until after Y is measured at $K+1$. As with A_K, we also take $A_{K,g}$ defined below to be 0. Furthermore we take R_0 to always be 0 since it is not the result of any test.) If the results of tests at time t were available immediately and hence could influence S_t, we would simply redefine R_t to be the results of testing at t rather than at $t-1$ and reorder as (L_t, A_t, R_t, S_t). Furthermore we let $\bar{H}_m \equiv (\bar{L}_m, \bar{R}_m, \bar{S}_{m-1}, \bar{A}_{m-1})$ be the past history through time m , excluding S_m, A_m. We let $\bar{\mathcal{H}}_m$ be the sample space of the random vector $\bar{H}_m$.

In Remark 5 of Section 8.1 we provide conditions under which the cost of testing c^* at time t can be made a function $c(t, \bar{H}_t, S_t)$ of the past information rather than a fixed constant.

A deterministic testing and treatment regime $g \equiv (g_0, g_1, \ldots, g_K)$ is a vector of rules or functions $g_t : (\bar{L}_t, \bar{R}_t, \bar{S}_{t-1}, \bar{A}_{t-1}) \mapsto (s_t, a_t)$ that determines the values (s_t, a_t) to which S_t and A_t will be set given the past $(\bar{L}_t, \bar{R}_t, \bar{S}_{t-1}, \bar{A}_{t-1})$. We denote arbitrary regimes by g and we adopt the counterfactual framework of Robins (1986, 1987) in which Y_g, Y_g^d, $L_{t,g}$, $R_{t,g}$, $S_{t,g}$, and $A_{t,g}$ are random variables representing the counterfactual data had regime g been followed. Implicit in the notation is the assumption that the treatment regime followed by one patient does not influence the outcome of any other patient. A testing and treatment regime is said to be *static* if g_t is a constant function for all t. A regime is said to be *dynamic* if it stipulates that testing and/or treatment at time t depends on past covariate and/or treatment values. A random testing and treatment regime replaces the functions g_t by conditional densities taking values in the support of (S_t, A_t). Since the optimal testing and treatment strategy can always be taken to be a deterministic strategy, we do not further consider random regimes in this article. We make three additional standard assumptions that serve to identify the optimal testing and treatment regime, even when the NDE assumption fails to hold (Robins 2004). Throughout, let

$$\Pi_m \equiv \pi_m(\bar{H}_m, S_m) \equiv \Pr[A_m = 1 | \bar{H}_m, S_m] \quad \text{and}$$
$$p(\cdot|\bar{H}_m) \equiv \Pr[S_m = \cdot | \bar{H}_m].$$

1. Positivity: for all $m = 0, \ldots, K$, Π_m and $p(s_m|\bar{H}_m)$ for all s_m in the sample space of S_m , are bounded away from 0 and 1 on a set of probability 1, except $\Pi_K = 0$;
2. Consistency: $Y^d = Y_g^d$ and $Y = Y_g$ if $(\bar{A}_{K,g}, \bar{S}_{K,g}) = (\bar{A}_K, \bar{S}_K)$, $(\bar{L}_{t+1}, \bar{R}_{t+1}, \bar{A}_{t+1}, \bar{S}_{t+1}) = (\bar{L}_{t+1,g}, \bar{R}_{t+1,g}, \bar{A}_{t+1,g}, \bar{S}_{t+1,g})$ if $(\bar{A}_{t,g}, \bar{S}_{t,g}) = (\bar{A}_t, \bar{S}_t)$;
3. Sequential exchangeability:

$$(Y_g^d, Y_g, \underline{L}_{t+1,g}, \underline{R}_{t+1,g}, \underline{A}_{t,g}, \underline{S}_{t,g}) \amalg (A_t, S_t)|\bar{L}_t, \bar{R}_t, (\bar{A}_{t-1}, \bar{S}_{t-1}) = \bar{g}_{t-1}(\bar{H}_{t-1}) \forall t, g,$$

where $\amalg$ stands for statistical independence and $\{(\bar{A}_{t-1}, \bar{S}_{t-1}) = \bar{g}_{t-1}(\bar{H}_{t-1})\} \equiv \{(A_m, S_m) = g_m(\bar{H}_m), m = 0, \ldots, t-1\}$.

Remark 1. The sample space of R_{t+1} includes "?" and R_{t+1} is always observed. When $A_t = 0$ (no screening at t), R_{t+1} is not missing; rather $R_{t+1} =$?. Note $R_{t+1} =$? if and only if $A_t = 0$. In Section 7, we introduce a new underlying variable R^*_{t+1} that is equal to R_{t+1} when $A_t = 1$ and is missing when $A_t = 0$. Until then the variable R^*_{t+1} is not needed as both treatment and screening decisions at time $t+1$ will depend only on always observed available information such as R_{t+1}.

Note that we could replace the vector $(Y_g^d, Y_g, \underline{L}_{t+1,g}, \underline{R}_{t+1,g}, \underline{A}_{t+1,g}, \underline{S}_{t+1,g})$ in Assumption 3 by $(Y_g^d, \underline{L}_{t+1,g}, \underline{R}_{t+1,g})$ because conditional on $(\bar{A}_{t-1}, \bar{S}_{t-1}) = \bar{g}_{t-1}(\bar{H}_{t-1})$, $(Y_g, \underline{A}_{t,g}, \underline{S}_{t,g})$ is a deterministic function of $(Y_g^d, \underline{L}_{t+1,g}, \underline{R}_{t+1,g})$ and $\bar{H}_t$. For future reference, we highlight that Assumption 3 implies

$$\mathrm{E}\left[Y_{\bar{S}_{t-1}, \bar{A}_{t-1}, s_t, a_t, \underline{g}_{t+1}} \middle| \bar{H}_t \right] = \mathrm{E}\left[Y_{\bar{S}_{t-1}, \bar{A}_{t-1}, s_t, a_t, \underline{g}_{t+1}} \middle| \bar{H}_t, A_t, S_t \right], \tag{1}$$

where for any g, $Y_{\bar{S}_{t-1}, \bar{A}_{t-1}, s_t, a_t, \underline{g}_{t+1}}$ denotes the counterfactual total utility Y when a subject takes her observed treatment history $(\bar{S}_{t-1}, \bar{A}_{t-1})$ through $t-1$, and possibly contrary to fact, takes (s_t, a_t) at t, and, for $t < K$ the subject follows the dynamic testing and treatment regime g from $t+1$ onward.

Throughout we call the set of three Assumptions 1–3 the identifying (ID) assumptions. Robins (1999) noted that Assumptions 2 and 3 do not impose restrictions on the observed data distribution. Robins (1999) showed that under the ID assumptions, for any (s_t, a_t), the conditional counterfactual mean $\mathrm{E}[Y_{\bar{S}_{t-1}, \bar{A}_{t-1}, \underline{g}_t}|\bar{H}_t = \bar{h}_t]$ is identified and it is equal to the following, so-called g-formula (Robins 1986)

$$\begin{aligned}
&\mathrm{E}\left[Y_{\bar{S}_{t-1}, \bar{A}_{t-1}, \underline{g}_t} \middle| \bar{H}_t = \bar{h}_t \right] \\
&= \int y f(y|\bar{h}_K, \bar{a}_K, \bar{s}_K) \prod_{m=t}^{K} I_{g_m(\bar{h}_m)}(a_m, s_m) \\
&\prod_{m=t+1}^{K} f(l_m, r_m|\bar{h}_{m-1}, \bar{a}_{m-1}, \bar{s}_{m-1}) d\underline{a}_m d\underline{s}_m d\underline{l}_{m+1} d\underline{r}_{m+1} dy.
\end{aligned}$$

Note that $\mathrm{E}[Y_{\bar{S}_{t-1}, \bar{A}_{t-1}, \underline{g}_t}|\bar{H}_t = \bar{h}_t]$ depends on the law of the observed data only through the conditional distributions $f(y|\bar{h}_K, a_K, s_K)$ and $f(l_m, r_m|\bar{h}_{m-1}, \bar{a}_{m-1}, \bar{s}_{m-1})$ for $m = t+1, \ldots, K$. Let $\mathcal{G}$ be the set of all regimes g satisfying the ID assumptions. Thus, for every $g \in \mathcal{G}$, $\mathrm{E}[Y_{\bar{S}_{t-1}, \bar{A}_{t-1}, \underline{g}_t}|\bar{H}_t = \bar{h}_t]$ and $\mathrm{E}[Y_g]$ are identified. Our goal is to estimate an optimal dynamic testing and treatment regime $g^{\mathrm{opt}} \equiv \arg\max_{g \in \mathcal{G}} \mathrm{E}[Y_g]$ and its corresponding value (i.e., expected utility) $\mathrm{E}[Y_{g^{\mathrm{opt}}}]$. Since we have excluded exceptional laws, g^{opt} is unique (Robins 2004).

Under the ID assumptions, g^{opt} can be computed by DP (Bellman 1952) as follows. Define $g^{\mathrm{opt}*} = (g_0^{\mathrm{opt}*}, \ldots, g_K^{\mathrm{opt}*})$ by the following backward recursion. First, we define $g_K^{\mathrm{opt}*}(\bar{H}_K) \equiv \arg\max_{s_K, a_K} \mathrm{E}[Y_{\bar{S}_{K-1}, \bar{A}_{K-1}, s_K, a_K}|\bar{H}_K]$ and recursively for $t = K-1, \ldots, 0$, we define $g_t^{\mathrm{opt}*}(\bar{H}_t) \equiv \arg\max_{s_t, a_t} \mathrm{E}[Y_{\bar{S}_{t-1}, \bar{A}_{t-1}, s_t, a_t, \underline{g}_{t+1}^{\mathrm{opt}*}}|\bar{H}_t]$. Robins (2004) proved that under the ID assumptions, $\mathrm{E}[Y_{\bar{S}_{t-1}, \bar{A}_{t-1}, \underline{g}_t^{\mathrm{opt}*}}|\bar{H}_t] = \max_{\underline{g}_t \in \underline{\mathcal{G}}_t} \mathrm{E}[Y_{\bar{S}_{t-1}, \bar{A}_{t-1}, \underline{g}_t}|\bar{H}_t]$, thus proving that the DP solution $g^{\mathrm{opt}*}$ agrees with the optimal treatment regime g^{opt} under the ID assumptions.

3. Estimator of opt-SNMM Without the NDE Assumption

In this section, we review the definition and the estimation of opt-SNMMs (Robins 2004) for estimating optimal dynamic regimes. For any regime g, define the testing and treatment effect contrast

$$\begin{aligned}
&\gamma_t^{\underline{g}_{t+1}}(\bar{H}_t, s_t, a_t) \\
&:= \mathrm{E}\Big[Y_{\bar{S}_{t-1}, \bar{A}_{t-1}, s_t, a_t, \underline{g}_{t+1}} - Y_{\bar{S}_{t-1}, \bar{A}_{t-1}, s_t = 0, a_t = 0, \underline{g}_{t+1}} \Big| \\
&\bar{H}_t, (S_t, A_t) = (s_t, a_t) \Big].
\end{aligned}$$

This contrast is the average causal effect, among subjects with the history $\bar{H}_t, (S_t, A_t) = (s_t, a_t)$, of setting possibly contrary to fact S_t and A_t both to 0 rather than to their observed values when, again possibly contrary to fact, the regime g is followed from time $t+1$ onward. Because $Y_{\bar{S}_{t-1}, \bar{A}_{t-1}, s_t = 0, a_t = 0, \underline{g}_{t+1}}$ does not depend on the free indices (s_t, a_t),

$$\begin{aligned}
&\arg\max_{(s_t, a_t)} \gamma_t^{\underline{g}_{t+1}}(\bar{H}_t, s_t, a_t) \\
&= \arg\max_{(s_t, a_t)} \mathrm{E}\left[Y_{\bar{S}_{t-1}, \bar{A}_{t-1}, s_t, a_t, \underline{g}_{t+1}} \middle| \bar{H}_t, (S_t, A_t) = (s_t, a_t) \right]
\end{aligned}$$

under the ID assumptions. It then follows, that by the arguments given in the preceding section, under the ID assumptions, the optimal testing and treatment regime g^{opt} is given by the following backward recursion: for $t = K, K-1, \ldots, 0$, $g_t^{\mathrm{opt}}(\bar{H}_t) := \arg\max_{(s_t, a_t)} \gamma_t^{\underline{g}_{t+1}^{\mathrm{opt}}}(\bar{H}_t, s_t, a_t)$. For $t = 0, \ldots, K$, define $\gamma_t(\bar{H}_t, s_t, a_t) \equiv \gamma_t^{\underline{g}_{t+1}^{\mathrm{opt}}}(\bar{H}_t, s_t, a_t)$ and $\left(S_t^{\mathrm{opt}}(\gamma_t), A_t^{\mathrm{opt}}(\gamma_t)\right) \equiv g_t^{\mathrm{opt}}(\bar{H}_t)$. Finally, define

$$\begin{aligned}
\Delta_t(\gamma_t; \underline{\gamma}_{t+1}) \equiv Y &- \gamma_t(\bar{H}_t, S_t, A_t) \\
&+ \sum_{m=t+1}^{K} \Big\{ \gamma_m\left(\bar{H}_m, S_m^{\mathrm{opt}}(\gamma_m), A_m^{\mathrm{opt}}(\gamma_m)\right) \\
&- \gamma_m\left(\bar{H}_m, S_m, A_m\right) \Big\},
\end{aligned} \tag{2}$$

where by convention $\sum_{m=K+1}^{K}(\cdot) \equiv 0$. By straightforward algebra it can be shown that for any t,

$$\mathrm{E}[\Delta_t(\gamma_t;\underline{\gamma}_{t+1})|\bar{H}_t,A_t,S_t] = \mathrm{E}[Y^{\mathrm{opt}}_{\bar{A}_{t-1},\bar{S}_{t-1},a_t=0,s_t=0,\underline{g}_{t+1}}|\bar{H}_t,A_t,S_t] \tag{3}$$

regardless of whether or not the sequential exchangeability holds. Heuristically $\Delta_t(\gamma_t;\underline{\gamma}_{t+1})$ mimics $Y^{\mathrm{opt}}_{\bar{A}_{t-1},\bar{S}_{t-1},a_t=0,s_t=0,\underline{g}_{t+1}}$ in the sense that they have the same mean given $(\bar{H}_t,A_t,S_t)$. Under the ID assumptions and Equation (3)

$$\mathrm{E}\left[\Delta_t(\gamma_t;\underline{\gamma}_{t+1})|\bar{H}_t,A_t,S_t\right] = \mathrm{E}\left[\Delta_t(\gamma_t;\underline{\gamma}_{t+1})|\bar{H}_t\right], \tag{4}$$

which implies that, for arbitrary functions $Q_t(s_t,a_t) \equiv q_t(\bar{H}_t,s_t,a_t)$ selected by the analyst, the random variable

$$\begin{aligned} U_t(q_t,\underline{\gamma}_t) \equiv &\left\{\Delta_t(\gamma_t;\underline{\gamma}_{t+1}) - \mathrm{E}[\Delta_t(\gamma_t;\underline{\gamma}_{t+1})|\bar{H}_t]\right\} \\ &\times\left\{Q_t(S_t,A_t) - \mathrm{E}[Q_t(S_t,A_t)|\bar{H}_t]\right\} \end{aligned} \tag{5}$$

has mean zero by Equation (4). In fact, Robins (2004) showed that $\gamma_t(\bar{H}_t,S_t,A_t)$ is the unique function of $(\bar{H}_t,S_t,A_t)$ satisfying $\gamma_t(\bar{H}_t,0,0)=0$ that also satisfies the condition $\mathrm{E}[U_t(q_t,\underline{\gamma}_t)] = 0$ for all q_t such that $U_t(q_t,\underline{\gamma}_t)^2$ has finite mean. Therefore, γ_t is identified through the system of equations (5) for all q_t. One could choose $Q_t(S_t,A_t)$ either to simplify the form or minimize the variance of the opt-SNMM estimator defined below.

An opt-SNMM assumes a parametric model for the treatment effect contrasts $\gamma_t(\bar{H}_t,s_t,a_t)$, that is, it assumes that

$$\gamma_t(\bar{H}_t,s_t,a_t) = \gamma_t(\bar{H}_t,s_t,a_t;\Psi_t^*), \tag{6}$$

where Ψ_t^* is an unknown parameter vector and $\gamma_t(\bar{H}_t,s_t,a_t;\Psi_t)$ is a known function equal to 0 whenever $s_t = a_t = 0$ or $\Psi_t = 0$.

Under the model the optimal testing and treatment choice at t is $(S_t^{\mathrm{opt}}(\Psi_t),A_t^{\mathrm{opt}}(\Psi_t)) \equiv \arg\max_{s_t,a_t}\gamma_t(\bar{H}_t,s_t,a_t;\Psi_t)$ for $\Psi=\Psi^*$. In an abuse of notation, we define $\Delta_t(\Psi_t;\underline{\Psi}_{t+1})$ just like $\Delta_t(\gamma_t;\underline{\gamma}_{t+1})$ but with the functions $\gamma_m(\bar{H}_m,S_m,A_m;\Psi_m)$ replacing the true functions $\gamma_m(\bar{H}_m,S_m,A_m)$.

The ID assumptions and the opt-SNMM model (6) determine a semiparametric model $\mathbb{M}_1$ for the observed data distribution defined by the restriction that there exists a unique $\Psi^* \equiv (\Psi_0^*,\ldots,\Psi_K^*)$ such that for all $t = 0,1,\ldots,K$, $\mathrm{E}[\Delta_t(\Psi_t^*;\underline{\Psi}_{t+1}^*)|\bar{H}_t,A_t,S_t] = \mathrm{E}[\Delta_t(\Psi_t^*;\underline{\Psi}_{t+1}^*)|\bar{H}_t]$. Before proceeding further we need to review some well-known results from semiparametric theory.

3.1. Review of Semiparametric Theory

For a law P in a semiparametric model $\mathbb{M}$, the tangent space $\Lambda \equiv \Lambda(P)$ at P is defined as the closed linear span of the scores $\phi(P)$ at P for all one-dimensional parametric submodels $\alpha \to P_\alpha$ such that $P_{\alpha=0}=P$. For a given functional of interest $\beta \equiv \beta(P)$, a random variable $\mathrm{IF} = \mathrm{IF}(P)$ is referred to as an influence function for β at P if $\mathrm{E}_P[\mathrm{IF}(P)] = 0$ and for each submodel P_α, $\partial\beta(P_\alpha)/\partial\alpha|_{\alpha=0} = \mathrm{E}_P[\mathrm{IF}(P)\phi(P)]$ for the score $\phi(P)$ of model P_α at $\alpha=0$. The efficient influence function $\mathrm{EIF} = \mathrm{EIF}(P)$ for a functional β is the unique influence function lying in Λ. A random variable IF is an influence function if and only if $\mathrm{IF} - \mathrm{EIF} \in \Lambda^\perp$ where $\Lambda^\perp$ is the ortho-complement to Λ in $L_{2,0}(P)$ where $L_{2,0}(P)$ is the subspace of $L_2(P)$ comprised of mean zero elements of $L_2(P)$. Hence the set $\mathcal{IF}$ of all influence functions (at P) is $\{\mathrm{EIF}+H; H\in\Lambda^\perp\}$. It follows that for any influence function IF, $\mathrm{EIF} = \mathrm{IF} - \Pi[\mathrm{IF}|\Lambda^\perp]$ with $\Pi = \Pi_P$ the projection operator in $L_{2,0}(P)$. The nuisance tangent space $\Lambda_{\mathrm{nuis}} \equiv \Lambda_{\mathrm{nuis}}(P)$ for β at P is the subspace of $\Lambda(P)$ generated by the scores of one-dimensional parametric submodels in $\mathbb{M}$ for which $\Psi(P_\alpha)$ is constant. The set $\mathcal{IF}$ of influence functions is contained in $\Lambda_{\mathrm{nuis}}^\perp$.

Estimators $\widehat{\beta}$ of $\beta=\beta(P)$ with the property that there exists a random variable $D_i \equiv d(O_i)$ with mean zero and finite variance under P such that $n^{1/2}(\widehat{\beta}-\beta) = n^{-1/2}\sum_{i=1}^n D_i + o_p(1)$ are called asymptotically linear at P. By the central limit theorem and Slutsky's theorem, any asymptotically linear estimator is consistent and asymptotically normal with asymptotic variance equal to $\mathrm{var}(D)$. Furthermore, any two asymptotically linear estimators, say $\widehat{\beta}_1$ and $\widehat{\beta}_2$ with the same D are asymptotically equivalent in the sense that $n^{1/2}(\widehat{\beta}_1-\widehat{\beta}_2) = o_p(1)$. An estimator of $\widehat{\beta}$ is regular at P in a semiparametric model $\mathbb{M}$ if its convergence to β is locally uniform (van der Vaart 1998). Any regular, asymptotically linear (RAL) estimator has D equal to some influence function IF. It follows that the minimal possible variance of any RAL estimator is $\mathrm{E}[\mathrm{EIF}^2]$, which is referred to as the semiparametric variance bound for β at P in model $\mathbb{M}$. If $\Lambda = \Lambda(P)$ is all of $L_{2,0}(P)$, then all RAL estimators have the same influence function at P.

3.2. Semiparametric Inference in Model $\mathbb{M}_1$

The set $\Lambda_{\mathrm{ancillary}}$ of conditional scores for $p[A_t,S_t|\bar{H}_t]$, for $t = 0,\cdots,K$ consists of all functions of $(A_t,S_t,\bar{H}_t)$ in $L_{2,0}(P)$ that have mean zero given $\bar{H}_t$. Let $\Lambda_{1,\mathrm{nuis}}$ denote the nuisance tangent space of model $\mathbb{M}_1$. Because $\Psi^* \equiv \Psi^*(P)$ does not depend on P through $p[A_t,S_t|\bar{H}_t]$ for $t=0,\cdots,K$, we conclude that $\Lambda_{\mathrm{ancillary}} \subset \Lambda_{1,\mathrm{nuis}}$.

Robins (2004, Theorem 3.3, eq. (3.10)) proved that under model $\mathbb{M}_1$,

$$\Lambda_{1,\mathrm{nuis}}^\perp = \left\{\mathbb{U}(\boldsymbol{q},\Psi^*) = \sum_{t=0}^{K} U_t(q_t,\underline{\Psi}_t^*); q_t(\bar{H}_t,a_t,s_t)\right\},$$

where the q_t are vector functions of $\dim(\Psi_t^*)$ that are unrestricted except for the requirement $\mathrm{E}[U_t(q_t)^2] < +\infty$ where $Q_t(s_t,a_t) \equiv q_t(\bar{H}_t,s_t,a_t)$ and

$$\begin{aligned} U_t(q_t,\underline{\Psi}_t) \equiv &\left\{\Delta_t(\Psi_t;\underline{\Psi}_{t+1}) - \mathrm{E}[\Delta_t(\Psi_t;\underline{\Psi}_{t+1})|\bar{H}_t]\right\} \\ &\times\left\{Q_t(S_t,A_t) - \mathrm{E}[Q_t(S_t,A_t)|\bar{H}_t]\right\}. \end{aligned}$$

Furthermore, $U_t(q_t,\underline{\Psi}_t)$ is a doubly robust estimating function in the sense that it still has mean zero at $\underline{\Psi}_t^*$ if either (but not both) $\mathrm{E}\left[\Delta_t(\Psi_t^*;\underline{\Psi}_{t+1}^*)|\bar{H}_t\right]$ or $\mathrm{E}[Q_t(S_t,A_t)|\bar{H}_t]$ is replaced by an arbitrary function of $\bar{H}_t$.

The identity $\mathrm{E}[\mathbb{U}(\boldsymbol{q},\Psi^*)] = 0$ suggests that one solve $\mathbb{P}_n[\mathbb{U}(\boldsymbol{q},\Psi)] = 0$ to estimate Ψ^*. Assuming that $\mathrm{E}[\mathbb{U}(\boldsymbol{q},\Psi)] = 0$ has a unique solution and that $\frac{\partial}{\partial\Psi}\mathrm{E}[\mathbb{U}(\boldsymbol{q},\Psi)]\big|_{\Psi=\Psi^*}$ is invertible, $\widehat{\Psi}(\boldsymbol{q})$ will be a RAL estimator of Ψ^* under standard regularity conditions. For a specific choice $\boldsymbol{q}_{\mathrm{opt}} = \boldsymbol{q}_{\mathrm{opt}}(P)$ of $\boldsymbol{q}$, the estimator $\widehat{\Psi}(\boldsymbol{q}_{\mathrm{opt}})$ attains semiparametric efficiency bound for regular estimators of Ψ^* under model $\mathbb{M}_1$. Because of its

dependence on P, $\boldsymbol{q}_{\text{opt}}$ would have to be estimated from the data. In this article, we decided not to consider the issue of finding $\boldsymbol{q}_{\text{opt}}$ and instead focus only on adjusting for the NDE assumption. Finding $\boldsymbol{q}_{\text{opt}}$ often leads to solving complicated integral equations. In practice, a heuristic choice of $\boldsymbol{q}$ that generally does not depend on the data can be used (Vansteelandt and Joffe 2014).

Notice that $\widehat{\Psi}(\boldsymbol{q})$ is infeasible because $\mathbb{U}(\boldsymbol{q}, \Psi^*)$ depends on the unknowns $\mathrm{E}[\Delta_t(\Psi_t; \underline{\Psi}_{t+1})|\bar{H}_t]$ and $\mathrm{E}[Q_t(S_t, A_t)|\bar{H}_t]$. A feasible estimator $\widetilde{\Psi}(\boldsymbol{q})$ solves $\mathbb{P}_n[\widehat{\mathbb{U}}(\boldsymbol{q}, \Psi)] = 0$ where $\widehat{\mathbb{U}}(\boldsymbol{q}, \Psi)$ is defined like $\mathbb{U}(\boldsymbol{q}, \Psi)$ but with estimates $\widehat{\mathrm{E}}[\Delta_t(\Psi_t; \underline{\Psi}_{t+1})|\bar{H}_t]$ and $\widehat{\mathrm{E}}[Q_t(S_t, A_t)|\bar{H}_t]$ replacing the unknown true expectations. We discuss feasible estimators in Section 6. Until then, to focus on important issues, we restrict our discussion to infeasible, also called oracle, estimators.

Given oracle estimators $\widehat{\Psi} \equiv \widehat{\Psi}(\boldsymbol{q})$, we can estimate the optimal value function $\theta_{g^{\text{opt}}} \equiv \mathrm{E}[Y_{g^{\text{opt}}}]$ by noting $\mathrm{E}[\Delta_0(\Psi_0^*; \underline{\Psi}_1^*)] = \mathrm{E}[Y_{a_0=0, \underline{g}_1^{\text{opt}}}]$ and hence $\mathrm{E}[Y_{g^{\text{opt}}}] = \mathrm{E}[\Delta_0(\Psi_0^*; \underline{\Psi}_1^*) + \gamma_0(L_0, R_0, S_0^{\text{opt}}(\underline{\Psi}_0^*), A_0^{\text{opt}}(\underline{\Psi}_0^*); \Psi_0^*)]$. Our oracle estimate of $\theta_{g^{\text{opt}}} = \mathrm{E}[Y_{g^{\text{opt}}}]$ is then $\widehat{\theta}_{\text{SNMM}, g^{\text{opt}}}$ that solves

$$\begin{aligned}
&\mathbb{P}_n[U_{\text{SNMM}, g^{\text{opt}}}(\theta, \widehat{\Psi})] = 0,\\
&U_{\text{SNMM}, g^{\text{opt}}}(\theta, \widehat{\Psi}) = \Delta_0(\widehat{\Psi}_0; \widehat{\underline{\Psi}}_1)\\
&\quad + \gamma_0(\bar{L}_0, \bar{R}_0, S_0^{\text{opt}}(\widehat{\underline{\Psi}}_0), A_0^{\text{opt}}(\widehat{\underline{\Psi}}_0); \widehat{\Psi}_0) - \theta. \qquad (7)
\end{aligned}$$

4. The NDE of Testing Assumption

The restriction

$$Y^d_{\bar{a}_K, \bar{s}_K} = Y^d_{\bar{a}'_K, \bar{s}_K} \; \forall \bar{a}_K, \bar{a}'_K, \bar{s}_K \qquad (8)$$

encodes the assumption that testing history $\bar{A}_K$ has no direct effect on observed health outcome Y^d not through treatment $\bar{S}_K$. We refer to Equation (8) as the NDE on Y^d [NDE(Y^d)] assumption. The assumption means that if we intervene and set the treatment history $\bar{S}_K$ to any $\bar{s}_K$, then further intervening on testing A at any time has no effect on the health outcome Y^d. (Note however that A_t has a direct effect on the total utility $Y = Y^d - c^* \sum_t A_t$ for $c^* \neq 0$ even under the NDE assumption for the health outcome Y^d.) We will write NDE as shorthand for NDE(Y^d).

Consider the model defined by the ID Assumptions 1–3 and the assumption (8). Robins (1999) showed that such a model determines a model $\mathbb{M}_{\text{NDE}}$ for the observed data characterized, possibly up to inequality constraints, by the following set of restrictions, with $W_t \equiv \prod_{m=t}^K p(S_m|\bar{H}_m)$

$$\mathrm{E}\left[\left.\frac{b_t(\bar{H}_t, \underline{S}_t, Y^d)}{W_{t+1}}\right| \bar{H}_t, S_t, A_t\right] = \mathrm{E}\left[\left.\frac{b_t(\bar{H}_t, \underline{S}_t, Y^d)}{W_{t+1}}\right| \bar{H}_t, S_t\right] \quad t = 0, \ldots, K \qquad (9)$$

for any function $b_t(\bar{H}_t, \underline{S}_t, Y^d)$ of all the history $\bar{H}_t$ up to time t, all the treatments $\underline{S}_t$ at t and onward, and the disease outcome Y^d.

The intuition behind these restrictions is that, under the three ID assumptions, weighting by W_{t+1}^{-1}, which is heuristically, the inverse of the probability of $\underline{S}_{t+1}$, mimics a randomized trial in which each of the treatments in $\underline{S}_{t+1}$ are assigned independently with probability 1/2 and the testing indicator A_t is randomly assigned given the past $(\bar{H}_t, S_t)$. Therefore, in this weighted distribution, A_t is independent of $(\underline{S}_{t+1}, Y^d)$ [and thus of any function $b_t(\bar{H}_t, \underline{S}_t, Y^d)$] conditional on $(\bar{H}_t, S_t)$, by $\underline{S}_{t+1}$ being completely randomized and the NDE assumption.

In what follows, for any $b_t, t = 0, \ldots, K$, we let $\mathcal{B}_t \equiv \left\{b_t(\bar{H}_t, \underline{S}_t, Y^d) : \mathrm{E}[T_{b,t}^2] < +\infty\right\}$ and $\mathcal{T}_t \equiv \left\{T_{b,t}; b_t \in \mathcal{B}_t\right\}$ where

$$\begin{aligned}
T_{b,t} \equiv D_{b,t} &- \left\{\mathrm{E}[D_{b,t}|\bar{H}_t, S_t, A_t] - \mathrm{E}[D_{b,t}|\bar{H}_t, S_t]\right\}\\
&- \sum_{m=t+1}^K \left\{\mathrm{E}[D_{b,t}|\bar{H}_m, S_m] - \mathrm{E}[D_{b,t}|\bar{H}_m]\right\}, \qquad (10)
\end{aligned}$$

$$\text{with } D_{b,t} \equiv \frac{b_t(\bar{H}_t, \underline{S}_t, Y^d)}{W_{t+1}}(A_t - \mathrm{E}[A_t|\bar{H}_t, S_t]). \qquad (11)$$

Here, $T_{b,t}$ is the residual from the projection of $D_{b,t}$ on the space

$$\Lambda_{\text{ancillary}} = \left\{V = \sum_{t=0}^K \left\{V_t - \mathrm{E}[V_t|\bar{H}_t]\right\}; \; V_t = v_t(\bar{H}_t, S_t, A_t) : \mathrm{E}[\mathrm{var}(V_t|\bar{H}_t)] < +\infty\right\}$$

of the conditional scores for the testing and treatment processes.

Remark 2. Under the ID assumptions, the NDE assumption is empirically falsifiable; since the NDE(Y^d) assumption implies $\mathrm{E}[D_{b,t}] = 0$ for all $b_t \in \mathcal{B}_t$. Owing to limited space, we do not consider falsification tests further.

Remark 3. The projection of any random variable $C = c(O)$ on conditional scores for testing and treatment is $\Pi[C|\Lambda_{\text{ancillary}}] = \sum_{t=0}^K (\mathrm{E}[C|\bar{H}_t, S_t, A_t] - \mathrm{E}[C|\bar{H}_t])$. In Appendix A.1, however, we show that the NDE assumption implies that if $C = D_{b,t}$, this projection does not depend on conditional scores for $\Pr(A_m = 1|\bar{H}_m, S_m)$ for $m > t$ and $\mathrm{E}[D_{b,t}|\bar{H}_t, S_t] = 0$ proving Equation (10) (for $m < t$, both $\mathrm{E}[C|\bar{H}_m, S_m, A_m]$ and $\mathrm{E}[C|\bar{H}_m]$ are zero by the NDE assumption).

We then have the following theorems, the proofs of which are deferred to Appendices A.1 and A.2.

Theorem 1. The space $\Lambda_{\text{NDE}}^{\perp}$ of mean zero random variables orthogonal to the tangent space Λ_{NDE} of model $\mathbb{M}_{\text{NDE}}$ is given by

$$\begin{aligned}
\Lambda_{\text{NDE}}^{\perp} &= \mathcal{T}_0 + \cdots + \mathcal{T}_K\\
&= \left\{T_b \equiv \sum_{t=0}^K T_{b,t}; b \equiv (b_0, \ldots, b_K) \quad \text{with} \right.\\
&\qquad \left. b_t \in \mathcal{B}_t, t = 0, \ldots, K\right\}.
\end{aligned}$$

Theorem 2. Each $T_b \in \Lambda_{\text{NDE}}^{\perp}$ is doubly robust in the sense that it has mean zero if either (i) the true densities $p(L_{k+1}|\bar{H}_k, S_k)$ are replaced by arbitrary densities $p^\dagger(L_{k+1}|\bar{H}_k, S_k)$ for $k = 0, \ldots, K$ except with $H_{K+1} \equiv Y^d$ replacing L_{K+1} or (ii) $p(S_k|\bar{H}_k)$ and Π_k are replaced by arbitrary conditional probability functions $p^\dagger(S_k|\bar{H}_k)$ and $\Pi_k^\dagger$ for $k = 0, \ldots, K$, but not both.

We prove Theorem 2 by actually proving a stronger version of double robustness in Appendix A.2, based on an alternative representation of any T_b.

5. Oracle Estimators of Ψ^* With Improved Efficiency

5.1. Suboptimal Estimators With Improved Efficiency

Let the semiparametric model $\mathbb{M}_{\text{int}} = \mathbb{M}_1 \cap \mathbb{M}_{\text{NDE}}$ consist of the set of observed data laws that satisfy the restrictions of models $\mathbb{M}_1$ and $\mathbb{M}_{\text{NDE}}$. Since under model $\mathbb{M}_{\text{NDE}}$, $\mathrm{E}[T_b] = 0$ for $b \equiv (b_0, \ldots, b_K)^\top$ where $T_b \equiv \sum_{t=0}^K T_{b,t}$ and $T_{b,t}$ is defined as in Equation (10), it follows that under model $\mathbb{M}_{\text{int}}$ we can consider the class of estimating functions $\left\{\mathbb{U}(\boldsymbol{q}, b, \Psi) = \mathbb{U}(\boldsymbol{q}, \Psi) - c_{\text{OLS}}(\boldsymbol{q}, b, \Psi) T_b; b\right\}$ where $c_{\text{OLS}}(\boldsymbol{q}, b, \Psi) = \mathrm{E}[\mathbb{U}(\boldsymbol{q}, \Psi) T_b^\top]\{\mathrm{E}[T_b T_b^\top]\}^{-1}$. Note that $\widehat{\Psi}(\boldsymbol{q})$ defined in Section 3 is the same as $\widehat{\Psi}(\boldsymbol{q}, b)$ for b identically zero. Define $J = \frac{\partial}{\partial \Psi^\top} \mathrm{E}[\mathbb{U}(\boldsymbol{q}, b)]\big|_{\Psi=\Psi^*} = \frac{\partial}{\partial \Psi^\top} \mathrm{E}[\mathbb{U}(\boldsymbol{q})]\big|_{\Psi=\Psi^*}$. Since J does not depend on b, we have the following result:

Theorem 3. Suppose $\widehat{\Psi}(\boldsymbol{q}, b)$ is a RAL estimator of Ψ^*. Then it has influence function $J^{-1}\{\mathbb{U}(\boldsymbol{q}, \Psi^*) - c_{\text{OLS}}(\boldsymbol{q}, b, \Psi) T_b\}$. Consequently, $\sqrt{n}\{\widehat{\Psi}(\boldsymbol{q}, b) - \Psi^*\}$ is asymptotically normal with mean zero and variance equal to

$$V_{\text{oracle}}(\boldsymbol{q}, b) \equiv J^{-1}(\mathrm{var}[\mathbb{U}(\boldsymbol{q}, \Psi)] - c_{\text{OLS}}(\boldsymbol{q}, b, \Psi) \mathrm{E}[T_b T_b^\top] c_{\text{OLS}}(\boldsymbol{q}, b, \Psi)^\top) J^{-\top}.$$

Furthermore, $\widehat{\Psi}(\boldsymbol{q})$ has influence function $J^{-1}\mathbb{U}(\boldsymbol{q}, \Psi^*)$ and asymptotic variance $V_{\text{oracle}}(\boldsymbol{q}, b = 0)$ greater than $V_{\text{oracle}}(\boldsymbol{q}, b)$ (in the positive definitive sense) whenever $\mathrm{E}[\mathbb{U}(\boldsymbol{q}, \Psi^*) T_b] \neq 0$.

5.2. Near Optimal Oracle Estimators With Improved Efficiency

Our goal is to improve efficiency by finding the function $b_{\text{opt}}(\boldsymbol{q})$ that minimizes $V_{\text{oracle}}(\boldsymbol{q}, b)$ for a given $\mathbf{q}$. To do so, we first define the orthogonal projection of a mean zero random variable U onto a closed linear space $\mathcal{F}$ of mean zero random variables to be the unique element $F^* \in \mathcal{F}$ defined as $F^* = \arg\min_{F \in \mathcal{F}} \mathrm{var}[U - F]$ in the positive definite sense. Now for any two nested subspaces $\Phi_l \subset \Phi_{l'} \subseteq \Lambda_{\text{NDE}}^\perp$, let T_{b_l} and $T_{b_{l'}}$ be equal to $\Pi[\mathbb{U}(\boldsymbol{q}, \Psi^*)|\Phi_l]$ and $\Pi[\mathbb{U}(\boldsymbol{q}, \Psi^*)|\Phi_{l'}]$. From the definition of a projection, $\mathrm{var}(T_{b_l}) \leq \mathrm{var}(T_{b_{l'}})$. It follows that $V_{\text{oracle}}(\boldsymbol{q}, b_l) \geq V_{\text{oracle}}(\boldsymbol{q}, b_{l'})$, demonstrating that the larger the subspace Φ of $\Lambda_{\text{NDE}}^\perp$ one projects on, the more efficient the estimator $\widehat{\Psi}(\boldsymbol{q}, b_{\text{sub}})$. It also follows that $b_{\text{opt}}(\boldsymbol{q})$ is the unique function $b = (b_0, \ldots, b_K)^\top$ that satisfies $T_b = \Pi[\mathbb{U}(\boldsymbol{q}, \Psi^*)|\Lambda_{\text{NDE}}^\perp]$. Here each $b_t \in \mathcal{B}_t$ is a column vector function of the same dimension as $\Psi_t^*, t = 0, \ldots, K$.

Our next task is to characterize $\Pi[U|\Lambda_{\text{NDE}}^\perp]$ for any random variable U. In Appendix A.3, we show that when U is a continuous random variable, $\Pi[U|\Lambda_{\text{NDE}}^\perp]$ is a solution to a set of complicated integral equations that do not exist in closed form. However, for U a discrete random variable with finite sample space we will give below a closed form expression for $\Pi[U|\Lambda_{\text{NDE}}^\perp]$. Furthermore, in the case of a continuous U, we will derive a closed form expression for $\Pi[U|\Omega]$ for a particular subspace $\Omega \subset \Lambda_{\text{NDE}}^\perp$. We will argue that the associated estimator $\widehat{\Psi}(\boldsymbol{q}, b_{\text{sub}}) \equiv \widehat{\Psi}(\boldsymbol{q}, b_{\text{sub}}(\boldsymbol{q}))$ where $T_{b_{\text{sub}}} = \Pi[\mathbb{U}(\boldsymbol{q}, \Psi^*)|\Omega]$ should have efficiency nearly equal to that of the computationally intractable optimal estimator $\widehat{\Psi}(\boldsymbol{q}, b_{\text{opt}}) \equiv \widehat{\Psi}(\boldsymbol{q}, b_{\text{opt}}(\boldsymbol{q}))$ with $T_{b_{\text{opt}}(\boldsymbol{q})} = \Pi[\mathbb{U}(\boldsymbol{q}, \Psi^*)|\Lambda_{\text{NDE}}^\perp]$. Our results are corollaries of the next theorem, proved in Appendix A.4.

In what follows for $t = 0, \ldots, K$, let T_t be a fixed $\delta_t \times 1$ random vector and let $\Gamma_t \equiv \left\{d_t(\bar{H}_t) T_t : d_t(\bar{H}_t) \text{ any } 1 \times \delta_t \text{ vector with } \mathrm{E}\left[d_t(\bar{H}_t)^2\right] < +\infty\right\}$. For $j = 0, \ldots, K$, define $T_j^{(K)} \equiv T_j$ and define also recursively for $t = K-1, \ldots, j$,

$$T_j^{(t)} \equiv T_j^{(t+1)} - \mathrm{E}\left[T_j^{(t+1)} T_{t+1}^{(t+1)\top} \middle| \bar{H}_{t+1}\right] \mathrm{E}\left[T_{t+1}^{(t+1)} T_{t+1}^{(t+1)\top} \middle| \bar{H}_{t+1}\right]^{-1} T_{t+1}^{(t+1)}.$$

Theorem 4. For any random variable U, $\Pi[U|\Gamma_0 + \ldots + \Gamma_K] = \sum_{t=0}^K d_t^*(\bar{H}_t) T_t$ where

$$d_0^*(\bar{H}_0) \equiv \mathrm{E}\left[U T_0^{(0)\top} \middle| \bar{H}_0\right] \mathrm{E}\left[T_0^{(0)} T_0^{(0)\top} \middle| \bar{H}_0\right]^{-1} \tag{12}$$

and for $t = 1, \ldots, K$,

$$d_t^*(\bar{H}_t) \equiv \mathrm{E}\left[\left\{U - \sum_{j=0}^{t-1} d_j^*(\bar{H}_j) T_j^{(t)}\right\} T_t^{(t)\top} \middle| \bar{H}_t\right] \mathrm{E}\left[T_t^{(t)} T_t^{(t)\top} \middle| \bar{H}_t\right]^{-1}. \tag{13}$$

The preceding theorem has the following important consequences. Let $\mathcal{S}_t$ denote the sample space of the treatment variable S_t and let I_t be the $\mathrm{card}(\mathcal{S}_t \times \cdots \times \mathcal{S}_K) \times 1$ vector whose elements are the indicators that $\underline{S}_t$ take a specific value $\underline{s}_t \in \mathcal{S}_t \times \cdots \times \mathcal{S}_K$, that is, $I_t \equiv (I_{\underline{s}_t}(\underline{S}_t))_{\underline{s}_t \in \mathcal{S}_t \times \cdots \times \mathcal{S}_K}$. Next, for any given $\xi \times 1$ vector $\boldsymbol{\varphi}(Y^d) \equiv (\varphi_1(Y^d), \ldots, \varphi_\xi(Y^d))^\top$ of linearly independent functions of Y^d, define for each $t = 0, \ldots, K$, the $\delta_t \times 1$ vector function

$$b_t^*(\bar{H}_t, \underline{S}_t, Y^d) = (\varphi_1(Y^d) I_t^\top, \varphi_2(Y^d) I_t^\top, \ldots, \varphi_\xi(Y^d) I_t^\top)^\top, \tag{14}$$

where $\delta_t \equiv \mathrm{card}(\mathcal{S}_t \times \cdots \times \mathcal{S}_K)\xi$ and the dependence on ξ has been suppressed in the notation. Note $b_t^*(\bar{H}_t, \underline{S}_t, Y^d)$ does not actually depend on $\bar{H}_t$. Define the set Ω_t to be the set Γ_t with $T_{b^*,t}$ (see Equation (10) for its form) substituted for T_t. Clearly Ω_t is a subspace of $\mathcal{T}_t$, defined in Theorem 1. In particular, we have the following important corollary.

Corollary 1. Consider the space $\Omega = \sum_{t=0}^K \Omega_t$. Then

$$\Pi[\mathbb{U}(\boldsymbol{q}, \Psi^*)|\Omega] = \sum_{t=0}^K d_t^*(\bar{H}_t) T_{b^*,t} =: \sum_{t=0}^K T_{b_{\text{sub}},t} \tag{15}$$

where $d_t^*(\bar{H}_t)$ is defined as in Equations (12) and (13) but with T_t replaced by $T_{b^*,t}$, and $b_{\text{sub}} = (b_{\text{sub},0}, \ldots, b_{\text{sub},K})^\top$, where the dependence of b_{sub} on $\mathbf{q}$ has been suppressed in the notation. In particular, when Y is discrete and ξ is the cardinality of the sample space of Y then $\Lambda_{\text{NDE}}^\perp = \Omega$ and consequently $\Pi[\mathbb{U}(\boldsymbol{q}, \Psi^*)|\Lambda_{\text{NDE}}^\perp] = \Pi[\mathbb{U}(\boldsymbol{q}, \Psi^*)|\Omega] = \sum_{t=0}^K T_{b_{\text{sub}},t}$.

Remark 4. Consider the case where Y is continuous and $\boldsymbol{\varphi}(Y)$ is the vector of the first ξ elements of a complete basis for $L_2(\mu)$ with μ the Lebesgue measure. Then as $\xi \to \infty$, $\Pi[\mathbb{U}(\boldsymbol{q}, \Psi^*)|\Omega] = \sum_{t=0}^{K} d_t^*(\bar{H}_t) T_{b^*,t}$ should converge to $\Pi[\mathbb{U}(\boldsymbol{q}, \Psi^*)|\Lambda_{\text{NDE}}^{\perp}]$. As a consequence, by choosing $\xi \to \infty$ slowly with the sample size n, the asymptotic variance $V_{\text{oracle}}(\boldsymbol{q}, b_{\text{sub}})$ of $\widehat{\Psi}(\boldsymbol{q}, b_{\text{sub}})$ should converge to the asymptotic variance $V_{\text{oracle}}(\boldsymbol{q}, b_{\text{opt}})$ of $\widehat{\Psi}(\boldsymbol{q}, b_{\text{opt}})$ and thus the oracle estimators $\widehat{\Psi}(\boldsymbol{q}, b_{\text{sub}})$ and $\widehat{\Psi}(\boldsymbol{q}, b_{\text{opt}})$ will be asymptotically equivalent.

6. Doubly Robust Nearly Efficient Estimation

In this section, we consider the construction of feasible estimators of Ψ^*. To do so, we shall need to estimate the unknown conditional means and densities (hereafter nuisance functions) that are present in $\widehat{\Psi}(\boldsymbol{q}, b)$, $\widehat{\Psi}(\boldsymbol{q}, b_{\text{sub}})$, $\widehat{\Psi}(\boldsymbol{q}, b_{\text{opt}})$ where the latter two estimators contain additional nuisance functions due to the dependence of b_{opt} and b_{sub} on the distribution P. In the following to cover all cases we write the generic oracle estimator as $\widehat{\Psi}(\boldsymbol{q}, b(P))$. We will consider the state-of-the-art cross-fit doubly robust machine learning (DR-ML) estimators (Chernozhukov et al. 2018, Smucler et al. 2019) $\widetilde{\Psi}_{cf}(\boldsymbol{q}, \widehat{b})$ in which the nuisance functions are estimated by arbitrary machine learning algorithms chosen by the analyst. [If $b(P) = b$ does not depend on P, then $\widehat{b} = b$.]

We need to use sample splitting because the functions estimated by modern machine learning algorithms (e.g., deep neural nets) have unknown statistical properties and, in particular, may not lie in a so-called Donsker class (see, e.g., van der Vaart and Wellner 1996, chap. 2)—a condition generally needed for asymptotic linearity when we do not split the sample. Cross-fitting can recover the information lost due to sample splitting, provided that $\widetilde{\Psi}_{cf}(\boldsymbol{q}, \widehat{b})$ is asymptotically linear.

The following algorithm computes $\widetilde{\Psi}_{cf}(\boldsymbol{q}, \widehat{b})$:

(i) The n study subjects are randomly split into two parts: an estimation sample of size n_1 and a nuisance sample of size $n_2 = n - n_1$ with $n_1/n \approx 1/2$.

(ii) Estimate all the unknown conditional expectation and density functions occurring in $\mathbb{U}(\boldsymbol{q}, b(P), \Psi) \equiv \mathbb{U}(\boldsymbol{q}, \Psi) - c_{\text{OLS}}(\boldsymbol{q}, b(P), \Psi)^\top T_{b(P)}$ from the nuisance sample data by machine learning. However, unconditional expectations can be estimated from the estimation sample.

(iii) Define $\widehat{\mathbb{U}}(\boldsymbol{q}, \widehat{b}, \Psi)$ to be $\mathbb{U}(\boldsymbol{q}, b(P), \Psi)$ except with the estimates under (ii) substituted for their estimands. Find the (assumed unique) solution $\widetilde{\Psi}^{(1)}(\boldsymbol{q}, \widehat{b})$ to $\mathbb{P}_n^{(1)}[\widehat{\mathbb{U}}(\boldsymbol{q}, \widehat{b}, \Psi)] = 0$ where $\mathbb{P}_n^{(1)}$ is the sample average operator in the estimation sample.

(iv) Let $\widetilde{\Psi}_{cf}(\boldsymbol{q}, \widehat{b}) = \left\{\widetilde{\Psi}^{(1)}(\boldsymbol{q}, \widehat{b}) + \widetilde{\Psi}^{(2)}(\boldsymbol{q}, \widehat{b})\right\}/2$ where $\widetilde{\Psi}^{(2)}(\boldsymbol{q}, \widehat{b})$ is $\widetilde{\Psi}^{(1)}(\boldsymbol{q}, \widehat{b})$ but with the *training* and *estimation* samples reversed.

In Appendix A.5, we provide regularity conditions under which $\widetilde{\Psi}_{cf}(\boldsymbol{q}, \widehat{b})$ is RAL with the same influence function as $\widehat{\Psi}(\boldsymbol{q}, b(P))$. Further we describe modified versions of $\widetilde{\Psi}_{cf}(\boldsymbol{q}, \widehat{b})$ that should exhibit better finite sample behavior. Finally, we show that under the linear model $\gamma_t(\bar{H}_t, S_t, A_t; \Psi_t) = \Psi_t^\top V_t$, for a given vector $V_t = v_t(\bar{H}_t, S_t, A_t)$, for each t, $\widetilde{\Psi}_{cf}(\boldsymbol{q}, b)$ exists in closed form and thus is simple to compute. A more detailed discussion of this strategy is given in Remark 10 in the Appendix and also see Kallus et al. (2019) for an extended discussion and theoretical details.

We let $\widetilde{\Psi}_{\text{ss}}(\boldsymbol{q}, \widehat{b})$ denote the single sample estimator solving $\mathbb{P}_n[\widehat{\mathbb{U}}(\boldsymbol{q}, \widehat{b}, \Psi)] = 0$ where $\mathbb{P}_n$ is the sample average operator in the entire sample and the unknown nuisance functions in $\mathbb{U}(\boldsymbol{q}, b(P), \Psi)$ are estimated from all n subjects. In the sequel, we will use $\widetilde{\Psi}$ to represent both estimators, as both are in common use, although the regularity conditions under which $\widetilde{\Psi}_{\text{ss}}(\boldsymbol{q}, \widehat{b})$ is RAL are more restrictive, generally requiring nuisance functions to lie in a Donsker class. For convenience, we will assume the necessary regularity conditions hold for any given $\widetilde{\Psi}$ to be RAL, as regularity conditions are not the focus of the article.

7. Varieties of NDE Assumptions—Substantive Plausibility

In this section, we first define additional NDE assumptions that are stronger than the NDE(Y^d) assumption and discuss their differing statistical implications and substantive meaning. Let $C_t \subseteq L_t, t = 0, \ldots, K$, be a (possibly improper or null) subset of the covariates L_t. A stronger assumption than (8) is that $\bar{A}_K$ has NDE on observed responses $\bar{C}_K$ and Y^d not through treatment $\bar{S}_K$. Formally, this is the assumption that

$$C_{t,\bar{a}_{t-1},\bar{s}_{t-1}} = C_{t,\bar{a}'_{t-1},\bar{s}_{t-1}}, Y^d_{\bar{a}_K,\bar{s}_K} = Y^d_{\bar{a}'_K,\bar{s}_K} \; \forall \bar{a}_K, \bar{a}'_K, \bar{s}_K, t, \quad (16)$$

which we refer to as the NDE($\bar{C}_K, Y^d$) assumption. Note, for any $C'_t \subset C_t$, the NDE($\bar{C}_K, Y^d$) assumption implies the NDE($\bar{C}'_K$), NDE($\bar{C}_K$) and NDE(Y^d) assumptions, where the NDE($\bar{C}_K$) assumption is given by Equation (16) without reference to Y^d. It follows from Robins (1999) that all the above results obtained under NDE(Y^d) hold under NDE($\bar{C}_K, Y^d$) when we replace $D_{b,t}$, $T_{b,t}$, $\mathcal{B}_t$, and $\Lambda_{\text{NDE}}^{\perp}$ with $D^c_{b,t}$, $T^c_{b,t}$, $\mathcal{B}^c_t$, and $\Lambda^{c\perp}_{\text{NDE}}$ where the latter are defined like the former except with the set of functions $\{b^c_t(\bar{H}_t, \underline{S}_t, \underline{C}_{t+1}, Y^d)\}$ replacing the set $\{b_t(\bar{H}_t, \underline{S}_t, Y^d)\}$. Since the latter set is contained in the former, it follows that the asymptotic variance $V_{\text{oracle}}(\boldsymbol{q}, b^c_{\text{opt}})$ of $\widehat{\Psi}(\boldsymbol{q}, b^c_{\text{opt}})$ is never greater and usually less than $V_{\text{oracle}}(\boldsymbol{q}, b_{\text{opt}})$ of $\widehat{\Psi}(\boldsymbol{q}, b_{\text{opt}})$, where b^c_{opt} is the minimizer of $V_{\text{oracle}}(\boldsymbol{q}, b^c)$ over functions in $\mathcal{B}^c_t$.

We next briefly discuss, via examples, how one might assess whether a particular NDE assumption is substantively plausible. Without thinking carefully, most tend to assume that a test will not have a direct effect on an outcome not through the treatment S_t under study. Therefore, it will be pedagogically most useful to give examples where this is not the case. In the HIV setting, a test may have an adverse effect on mortality Y^d not through ART S_t, hence falsifying the NDE (Y^d) assumption, if poor test results (high HIV RNA and low CD4 count) motivate physicians to provide life-extending ancillary care such as prophylactic therapy against opportunistic infections or a list of avoidable behaviors likely to increase viral replication. This is true whether the ancillary care is generally (prophylaxis therapy) or rarely (a list of avoidable behaviors) recorded in the medical record.

Consider next an example in which a test has no effect on Y^d except through S_t but does have an effect on a covariate C_t. A cystoscopic exam A_{t-1} to evaluate recurrence of a bladder cancer can cause intense pain $C_t = 1$ in certain individuals. Pain itself has no effect on bladder cancer mortality Y^d, except by (possibly delaying) chemotherapy S_t; hence NDE(Y^d) is true but NDE($\bar{C}_K, Y^d$) is false. In this setting, an analyst who uses the estimator $\widehat{\Psi}(\boldsymbol{q}, b_{\text{opt}})$ will succeed in gaining efficiency without incurring bias while an analyst who uses $\widehat{\Psi}(\boldsymbol{q}, b^c_{\text{opt}})$ may introduce bias as T^c_b may not have mean zero.

Finally, we consider an example wherein the assumption that "test at time t has NDE on the results of later tests" is false. To formalize, we follow Robins, Orellana, and Rotnitzky (2008) and assume the existence of an additional variable R^*_t that in conjunction with A_{t-1} determines R_t by $R_t = A_{t-1}R^*_t + (1 - A_{t-1})?$. That is, if a test is performed at $t-1$, the underlying variable R^*_t is revealed; otherwise R^*_t is hidden and recorded as ?. Let $R^*_{t,\bar{a}_{t-1},\bar{s}_{t-1}}$ and $R^*_{t,g}$ be the corresponding counterfactuals. We assume R^*_t just precedes R_t in our temporal ordering, as would be the case if a test (e.g. a cardiac stress test) ordered at visit $t-1$ was both conducted and results reported at visit t. (With minor notational changes, we could have defined $R_t = A_{t-1}R^*_{t-1} + (1 - A_{t-1})?$, with R^*_{t-1} just preceding S_{t-1}, if the results of a test ordered and conducted at $t-1$ are not available till t.) We say that the NDE($\bar{R}^*_K, Y^d$) assumption of no effect of testing on $\bar{R}^*_K, Y^d$ holds if

$$R^*_{t,\bar{a}_{t-1},\bar{s}_{t-1}} = R^*_{t,\bar{a}_{t-1}\bar{s}_{t-1}}, Y^d_{\bar{a}_K,\bar{s}_K} = Y^d_{\bar{a}_K\bar{s}_K}, \forall \bar{a}_t, \bar{a}'_t, \bar{s}_{t-1}, t.$$

Consider a study of an effect of a drug S_t on brain amyloid content Y^d in early Alzheimer's disease in which a test A_t of mental ability R^*_{t+1} is repeatedly administered to some but not other participants. Then, even when the NDE(Y^d) holds, the NDE($\bar{R}^*_K, Y^d$) assumption will not hold if repetition has a direct effect on later test scores R^*_t due to "practice effects." That is, $R^*_{t,\bar{a}_{t-1}=\bar{0}_{t-1},a_t=1,\bar{s}_{t-1}}$ will generally be less than $R^*_{t,\bar{a}_{t-1}=\bar{1}_{t-1},a_t=1,\bar{s}_{t-1}}$.

Furthermore as in Robins (1999) we modify the identifying assumptions to incorporate $\bar{R}^*_K$, by (a) modifying the sequential exchangeability assumption both by adding the counterfactual $\underline{R}_{t+1,g}$ and conditioning on $\bar{R}^*_t$ (see equation (27) in Appendix A.7), (b) having $\bar{R}^*_K$ satisfy consistency, and (c) assuming $\bar{R}^*_t \amalg (S_t, A_t)|\bar{H}_t$ to insure that the modified ID assumptions both imply the unmodified ID assumptions and leave $\Pi_m \equiv \Pr[A_m = 1|\bar{H}_m, S_m]$ and $p(s_m|\bar{H}_m)$ unchanged, with no dependence on $\bar{R}^*_m$ except through $\bar{R}_m$.

Suppose the NDE($\bar{R}^*_K, Y^d$) and modified ID assumptions hold. Then we would like to use the estimator $\widehat{\Psi}(\boldsymbol{q}, b^{r^*}_{\text{opt}})$ where $T^r_{b^{r^*}_{\text{opt}}} \equiv \Pi[\mathbb{U}(\boldsymbol{q}, \Psi^*)|\Lambda^{r^*\perp}_{\text{NDE}}]$. However it is not possible to project onto $\Lambda^{r^*\perp}_{\text{NDE}}$, because though $\underline{R}_{t+1}$ is observed, $\underline{R}^*_{t+1}$ is unobserved except for subjects with $\underline{A}_{t+1} = \underline{1}_{t+1}$. Instead, we project onto the ortho-complement of the tangent space of the induced observed data model. Specifically and, more generally, let $\mathbb{M}_{\text{NDE}}(\bar{C}_K, \bar{R}^*_K, Y^d)$ be the model defined by the NDE($\bar{R}^*_K, \bar{C}_K, Y^d$) and modified ID assumptions. Let $\mathbb{M}_{\text{NDE,obs}}(\bar{C}_K, \bar{R}^*_K, Y^d)$ be the induced model for the observed data O. Let $\Lambda^{c,r^*\perp}_{\text{NDE,obs}}$ be the observed data ortho-complement to the tangent space of $\mathbb{M}_{\text{NDE,obs}}(\bar{C}_K, \bar{R}^*_K, Y^d)$. In Appendix A.7 we show that $\Lambda^{c,r^*\perp}_{\text{NDE,obs}}$ is a strict subset of $\Lambda^{c,r^*\perp}_{\text{NDE}}$, but a superset of $\Lambda^{c\perp}_{\text{NDE}}$. In Theorem 6 and Corollary 4 of Appendix A.7, we provide closed-form expressions both for $\Lambda^{c,r^*\perp}_{\text{NDE,obs}}$ and for the projection of any random variable onto a large subspace $\Omega^{c,r^*}_{\text{obs}}$ of $\Lambda^{c,r^*\perp}_{\text{NDE,obs}}$ using Theorem 4 and Corollary 1. Hence to gain maximum efficiency we use the algorithm of Section 6 to construct feasible versions of the oracle estimators $\widehat{\Psi}(\boldsymbol{q}, b^{c,r^*}_{\text{opt,obs}})$ or $\widehat{\Psi}(\boldsymbol{q}, b^{c,r^*}_{\text{sub,obs}})$ where $T_{b^{c,r^*}_{\text{opt}},\text{obs}} = \Pi[\mathbb{U}(\boldsymbol{q}, \Psi^*)|\Lambda^{c,r^*\perp}_{\text{NDE,obs}}]$ and $T_{b^{c,r^*}_{\text{sub}},\text{obs}} = \Pi[\mathbb{U}(\boldsymbol{q}, \Psi^*)|\Omega^{c,r^*}_{\text{obs}}]$.

8. Efficiency and Intuition

Above we described how to construct novel highly efficient estimators under various NDE assumptions. In this section, we provide guidance and intuition as to when novel estimators will achieve small versus large gains in efficiency compared to estimators that do not incorporate any NDE assumption. To do so, we begin by discussing the article of Caniglia et al. (2019). The authors of that article analyzed data from an observational study that followed 41,724 HIV infected individuals on first line anti-retroviral therapy (ART) whose HIV RNA had been suppressed to less than 200 copies/ml at baseline. They estimated the optimal testing and treatment strategy among four candidate regimes. Testing costs were ignored. The utility $Y = Y^d$ was the indicator of survival at 60 months of follow-up; hence the expected counterfactual utility $\theta_g \equiv \text{E}[Y_g]$ is the probability of surviving to 60 months under regime g. See Robins, Orellana, and Rotnitzky (2008) and later Remark 5 for the case of costly tests. Data were recorded at monthly intervals and the ID assumptions of Section 2 were assumed to hold. Their analysis compared the standard IPW estimator $\widetilde{\theta}_{\text{ipw},g}$ with a novel IPW estimator $\widetilde{\theta}_{\text{nde-ipw},g}$ introduced in Robins, Orellana, and Rotnitzky (2008). Below we show that the latter is a consistent estimator of θ_g under the NDE($\bar{R}^*_K, Y^d$) assumption, but not under the weaker NDE(Y^d) assumption; the former is consistent regardless of whether any NDE assumption holds. They found $\widetilde{\theta}_{\text{nde-ipw},g}$ to be 50 times as efficient as the usual IPW estimator $\widetilde{\theta}_{\text{ipw},g}$. In this subsection we describe results in Caniglia et al. (2019) and provide an intuitive understanding of the reason for such a remarkable efficiency gain.

Caniglia et al. (2019) compared θ_g for four candidate regimes g formed by crossing two testing with two treatment regimes. For reasons of both pedagogy and space, in the main text, we describe a somewhat simplified version of these regimes. Appendix A.6 compares these approximate regimes with the actual regimes used for data analysis. The two testing regimes g_x require simultaneous tests for both HIV RNA and CD4 count at baseline and then every 11 months while RNA < 200 copies/ml and CD4 $> x$ counts/ml for $x \in \{350, 500\}$; otherwise the tests are performed every 5 months. The two treatment regimes g_z require one to switch to second line ART therapy 2 months after the first time the measured RNA level exceeds z copies for $z \in \{200, 1000\}$. Let $g_{x,z} = (g_x, g_z), x \in \{350, 500\}, z \in \{200, 1000\}$ denote the four regimes.

At baseline all 41,724 subjects can follow each of the 4 regimes. To avoid unnecessary clutter, we henceforth restrict attention to a comparison of the regimes $g_{x=350,z=200}$ and $g_{x=500,z=200}$. Consider first an analysis that does not impose the NDE assumption. A subject is censored under (i.e., stops following) regime $g_{x,z}$ at the first month t that either the subject's observed treatment indicator S_t or testing indicator A_t is incompatible with $g_{x,z}$, meaning that either $S_{t,g} \neq S_t$ or $A_{t,g} \neq A_t$. For regime $g_{x=500,z=200}$, of the 41,724 initially following the regime, the number still following the regime (i.e., still uncensored) at month 12, 24, 36, 48, 60 were 8926, 2634, 949, 348, 152. The corresponding numbers 4806, 1006, 204, 94, 45 for $g_{x=350,z=200}$ were even smaller. The explanation for the rapid decrease in the number of uncensored subjects is that most individuals were tested more frequently in the observed data than is allowed by either of the two testing regimes under consideration. In fact, even under regime $g_{x=500,z=200}$, well over 95% of the censoring events were due to early retest (i.e., $A_{t,g} = 0, A_t = 1$).

Caniglia et al. (2019) constructed a so-called NDE dataset in which subjects are no longer censored from regime $g_{x,z}$ for reasons of overly frequent testing, thereby increasing the number of uncensored subjects by roughly one to two orders of magnitude. The estimator $\widetilde{\theta}_{\text{ipw-nde},g_{x,z}}$ is computed from the uncensored subjects in the NDE dataset and thus would be expected to have a much smaller variance than $\widetilde{\theta}_{\text{ipw},g_{x,z}}$. In fact, Caniglia et al. (2019) reported the nonparametric bootstrap variance estimators of the two different IPW estimators of the difference $\theta_{g_{x=350,z=200}} - \theta_{g_{x=500,z=200}}$. The nonparametric bootstrap estimator of the variance of $\widetilde{\theta}_{\text{ipw},g_{x=350,z=200}} - \widetilde{\theta}_{\text{ipw},g_{x=500,z=200}}$ was 1.69×10^{-4} compared to 3.22×10^{-6} for the variance of $\widetilde{\theta}_{\text{ipw-nde},g_{x=350,z=200}} - \widetilde{\theta}_{\text{ipw-nde},g_{x=500,z=200}}$, indicating a relative efficiency in favor of the latter of 53-fold. However, as noted above, if the relevant NDE assumptions fail to hold, $\widetilde{\theta}_{\text{ipw-nde},g_{x,z}}$ is inconsistent for $\theta_{g_{x,z}}$.

8.1. Definitions and Statistical Properties of $\widetilde{\theta}_{\text{ipw},g}$ and $\widetilde{\theta}_{\text{nde-ipw},g}$

Prior to the current article the estimator $\widetilde{\theta}_{\text{nde-ipw},g}$ was the only estimator in the literature that leveraged the NDE assumptions to increase efficiency. Thus, it is of interest to understand how $\widetilde{\theta}_{\text{nde-ipw},g}$ is related to the estimators proposed in the current article that leverage NDE assumptions by subtracting from an inefficient estimator the projection of its influence function onto the ortho-complement to the tangent space of an NDE model. In this section, we study this relationship after we formally define $\widetilde{\theta}_{\text{nde-ipw},g}$. Recall $W_t \equiv \prod_{m=t}^{K} p(S_m|\bar{H}_m)$ and $\Pi_m \equiv \Pr(A_m = 1|\bar{H}_m, S_m)$ and let $\widehat{W}_t$ and $\widehat{\Pi}_m$ be corresponding estimates. We ignore testing costs so $Y = Y^d$. Under the ID assumptions and regardless of whether the NDE assumption holds, $\theta_g = \mathrm{E}[Y_g]$ is identified by the IPW representation $\mathrm{E}[V_{\text{ipw},g}]$ of the g-formula where

$$V_{\text{ipw},g} \equiv \frac{\prod_{t=0}^{K} \mathbb{1}\{A_t = A_{t,g}, S_t = S_{t,g}\} Y}{\prod_{t=0}^{K} \Pr(A_t = A_{t,g}|\bar{H}_t, S_t) W_0}$$

$$= \prod_{t=0}^{K} \left\{ \frac{\mathbb{1}\{A_t = 0\}}{1 - \Pi_t} \right\}^{1-A_{t,g}} \prod_{t=0}^{K-1} \left\{ \frac{\mathbb{1}\{A_t = 1\}}{\Pi_t} \right\}^{A_{t,g}} \times \frac{\prod_{t=0}^{K} \mathbb{1}\{S_t = S_{t,g}\}}{W_0} Y.$$

Note the $K-1$ in the last expression is licensed by our earlier assumption that $A_K = A_{K,g} = \Pi_K = 0$. The quantities $\widehat{\theta}_{\text{ipw},g} = \mathbb{P}_n[V_{\text{ipw},g}]$ and $\widetilde{\theta}_{\text{ipw},g} = \mathbb{P}_n[\widehat{V}_{\text{ipw},g}]$ are, respectively, the oracle IPW estimator and the IPW estimator of θ_g , where, for any "oracle" random vector Q depending on unknown conditional expectations and densities, $\widehat{Q}$ replaces these unknown functions by estimates. By consistency and $g = (g_a, g_s)$ having both the testing and treatment regimes g_a and g_s deterministic, one can easily show that $V_{\text{ipw},g}$ is a function of the observed data assuming the oracle knowledge that all the nuisance functions are known. (Hence $\widehat{V}_{\text{ipw},g}$ is a statistic without such oracle knowledge.) Furthermore, a necessary condition for $V_{\text{ipw},g}$ to be nonzero is that $\bar{H}_{t,g} = \bar{H}_t$ for $t = 0, \ldots, K$, and $S_K = g_{s,K}(\bar{H}_K)$. A subject is said to have been censored under regime g if and only if for some $t \leq K$ at least one of the following 4 events happens: $[S_t = 0, S_{t,g} = 1]$, $[S_t = 1, S_{t,g} = 0]$, $[A_t = 0, A_{t,g} = 1]$, and $[A_t = 1, A_{t,g} = 0]$. Hence all censored subjects have $V_{\text{ipw},g} = 0$.

Robins, Orellana, and Rotnitzky (2008) introduced the NDE-IPW estimator $\widetilde{\theta}_{\text{nde-ipw},g} = \mathbb{P}_n[\widehat{V}_{\text{nde-ipw},g}]$ with oracle version $\widehat{\theta}_{\text{nde-ipw},g} = \mathbb{P}_n[V_{\text{nde-ipw},g}]$, where

$$V_{\text{nde-ipw},g} = \prod_{t=0}^{K-1} \left\{ \frac{\mathbb{1}\{A_t = 1\}}{\Pi_t} \right\}^{A_{t,g}} \frac{\prod_{t=0}^{K} \mathbb{1}\{S_t = S_{t,g}\}}{W_0} Y.$$

Note $V_{\text{nde-ipw},g}$ differs from $V_{\text{ipw},g}$ only in that we have removed the first factor in set braces. A subject who experiences the event $[A_t = 1, A_{t,g} = 0]$ need not have $V_{\text{nde-ipw},g} = 0$. Therefore, we will say a subject is NDE-censored under g (and thus not included in the NDE dataset under regime g) if and only if, for some $t \leq K$, at least one of the three NDE censoring events $[S_t = 0, S_{t,g} = 1]$, $[S_t = 1, S_{t,g} = 0]$, and $[A_t = 0, A_{t,g} = 1]$ happens.

Remark 5. If we wish to incorporate testing costs so Y need not equal Y^d, it is necessary to replace Y in the definition of $V_{\text{nde-ipw},g}$ by $Y_g = Y + \sum_{t=0}^{K-1} c^* \mathbb{1}[A_t = 1, A_{t,g} = 0]$ where Y_g adds to Y the number of times t that a subject experiences the event $[A_t = 1, A_{t,g} = 0]$ multiplied by the cost c^* of the test. To see why, note that, for non-NDE-censored subjects under regime g, the observed total utility Y is less than the utility Y_g under regime g by c^* times the difference in the number of tests $\sum_{t=0}^{K-1} \mathbb{1}[A_t = 1, A_{t,g} = 0]$. All the results given below continue to hold when testing costs are included if we substitute Y_g for Y in the definitions of $V_{\text{nde-ipw},g}$ and $V_{\text{ipw},g}$.

There is a subtle, but critical point we have glossed over: assuming that the nuisance functions are known, it is not immediately clear that $V_{\text{nde-ipw},g}$ itself is a statistic (i.e., a function

of the observed data O with probability 1) because, unlike for $V_{\text{ipw},g}$, an appeal to consistency does not suffice (since if the event $[A_{t-1} = 1, A_{t-1,g} = 0]$ occurs, then $\bar{H}_{t',g} \neq \bar{H}_{t'}$ for $t' \geq t$) and $\bar{H}_{t',g}$ may then not be a function of the observed data. Given a testing and treatment regime g of interest, the following lemma, proved in Appendix A.8, provides sufficient conditions for $V_{\text{nde-ipw},g}$ to be a statistic. As in Section 7, we assume there exists an underlying variable R^*_{t+1} that is revealed only if a test is performed, that is, $R_{t+1} = A_t R^*_{t+1} + (1 - A_t)$?.

Lemma 1. Suppose for regime g (i) $g_t(\bar{h}_t)$ depends on $\bar{h}_t$ through and only through the testing results $\bar{r}_t$, covariates $\bar{c}_t$, and treatments $\bar{s}_{t-1}$; equivalently $g_t(\bar{h}_t) = g^\dagger_t(\bar{h}^\dagger_t)$ for known $g^\dagger_t$ where $\bar{h}^\dagger_t = (\bar{s}_{t-1}, \bar{c}_t, \bar{r}_t)$. Then, if (ii) the NDE$(\bar{C}_K, \bar{R}^*_K)$ assumption holds, $V_{\text{nde-ipw},g}$ is a statistic.

Remark 6.

1. We could have included $\bar{a}_{t-1}$ in $\bar{h}^\dagger_{t-1}$ but it would have been redundant since $\bar{a}_{t-1}$ is deterministic function of $\bar{r}_t$, as $r_m =$? implies $a_{m-1} = 0$ and $r_m \neq$? implies $a_{m-1} = 1$ for $1 \leq m \leq K$. Had we included $\bar{a}_{t-1}$ in $\bar{h}^\dagger_{t-1}$, then $\bar{h}_t = (\bar{h}^\dagger_t, \bar{l}_{t-1} \setminus \bar{c}_t)$ would differ from $\bar{h}^\dagger_t$ only through the components $\bar{l}_t \setminus \bar{c}_t$ of $\bar{l}_t$ not included in $\bar{c}_t$. Note further the Lemma implies that, if the NDE$(\bar{C}_K, \bar{R}^*_K)$ assumption holds, then, for non-NDE-censored subjects, $Y_g = Y + \sum_{t=0}^{K-1} \mathbb{1}[A_t = 1, A_{t,g} = 0] c^*(t, \bar{C}_t, \bar{R}_t, \bar{S}_t)$ for cost function $c^*(t, \bar{c}_t, \bar{r}_t, \bar{s}_t)$ is a statistic and hence generalizes Remark 5 by allowing $c^*(t, \bar{C}_t, \bar{R}_t, \bar{S}_t)$ to replace c^*. See Remark 14 in Appendix A.8 for additional discussion.
2. To understand the need for the NDE$(\bar{R}^*_K)$ assumption in Lemma 1, suppose the assumption fails to hold and consider the following counterexample. Consider a regime g defined by $S_{0,g} = 0$, $A_{0,g} = 0$, $S_{1,g} = 1$, $A_{1,g} = 1$, $S_{2,g} = g_{s,2}(R_{2,g})$. Then, for a subject with observed data $S_0 = 0$, $A_0 = 1$, $S_1 = 1$, $A_1 = 1$, $R_2 \equiv R^*_{2,a_0=1,a_1=1,s_0=0,s_1=1} = 1$, $S_2 = 1$, we cannot compute $R_{2,g} \equiv R^*_{2,a_0=0,a_1=1,s_0=0,s_1=1}$ and thus $S_{2,g}$ from the observed data because $A_0 \neq A_{0,g}$. Hence $\widehat{V}_{\text{nde-ipw},g}$ is not a statistic. However if the NDE$(\bar{R}^*_K)$ assumption holds, $R_{2,g} = R_2 = R^*_{2,s_0=0,s_1=1}$ and thus $\widehat{V}_{\text{nde-ipw},g}$ is a statistic. An analogous counterexample can be constructed when the NDE$(\bar{C}_K)$ assumption fails to hold.

The next lemma gives conditions under which $V_{\text{nde-ipw},g}$ is unbiased for $\theta_g = \mathrm{E}[Y_g]$.

Lemma 2. Suppose the modified ID assumptions of Section 7 hold and g_t depends on and only on $\bar{r}_t$, $\bar{c}_t$, and $\bar{s}_{t-1}$. Then the NDE$(\bar{C}_K, \bar{R}^*_K, Y^d)$ assumption implies that $\mathrm{E}[V_{\text{nde-ipw},g}] = \theta_g$ and, thus, also that $\widetilde{\theta}_{\text{nde-ipw},g} = \mathbb{P}_n[\widehat{V}_{\text{nde-ipw},g}]$ is a RAL estimator of θ_g under regularity conditions.

Remark 7. It is of interest that the NDE$(\bar{C}_K, \bar{R}^*_K)$ assumption was already required for $\widehat{V}_{\text{nde-ipw},g}$ to even be a statistic. Robins, Orellana, and Rotnitzky (2008) proved Lemma 2 under the stronger NDE$(\bar{L}_K, \bar{R}^*_K, Y^d)$ assumption; inspection of their proof shows they actually proved Lemma 2; in fact, they showed that Lemma 2 remains true without requiring $\Pi_m \equiv \Pr(A_m = 1|\bar{H}_m, S_m)$ be bounded away from 1.

Here we offer an alternate proof of this lemma that reveals the relationship of $V_{\text{nde-ipw},g}$ to our earlier discussions of semiparametric models and their tangent spaces. Corollary 2 below shows that Lemma 2 follows directly from Lemma 3. In our proof of Lemma 3 in Appendix A.8, we show that the estimator $\widehat{\theta}_{\text{nde-ipw},g}$ (and its feasible version $\widetilde{\theta}_{\text{nde-ipw},g}$) is a member of a much larger class of estimators to which the results we obtain for $\widehat{\theta}_{\text{nde-ipw},g}$ (and its feasible version $\widetilde{\theta}_{\text{nde-ipw},g}$) also apply.

The model $\mathbb{M}_{np,\text{obs}}$ for the observed data defined by the modified ID assumptions has tangent space $\Lambda = L_{2,0}(P)$, as the model places no restrictions on the distribution of O. It follows that the functional $\theta_g = \mathrm{E}[V_{\text{ipw},g}]$ has a unique influence function $\text{IF}_{\text{ipw},g} \equiv V_{\text{ipw},g} - \Pi[V_{\text{ipw},g}|\Lambda_{\text{ancillary}}] - \theta_g$ in model $\mathbb{M}_{np,\text{obs}}$, which is thus also an influence function in the smaller model $\mathbb{M}_{\text{NDE,obs}}(\bar{C}_K, \bar{R}^*_K, Y^d)$.

Lemma 3. Under the assumptions of Lemma 2, $\text{IF}_{\text{ipw},g} - \theta_g - \left\{V_{\text{nde-ipw},g} - \Pi[V_{\text{nde-ipw},g}|\Lambda_{\text{ancillary}}]\right\} \in \Lambda^{c,r^*\perp}_{\text{NDE,obs}}$.

The corollary below then follows from the fact that in any semiparametric model the sum of any element of the orthocomplement to the tangent space and an influence function is itself an influence function.

Corollary 2. Under the assumptions of Lemma 2, $\text{IF}^{c,r^*}_{\text{nde-ipw},g} \equiv V_{\text{nde-ipw},g} - \Pi[V_{\text{nde-ipw},g}|\Lambda_{\text{ancillary}}] - \theta_g$ is an influence function in model $\mathbb{M}_{\text{NDE,obs}}(\bar{C}_K, \bar{R}^*_K, Y^d)$ and thus $V_{\text{nde-ipw},g}$ has mean θ_g.

This following remark summarizes the implications of the above results.

Remark 8.

1. Lemma 3 fails to hold if we replace $\Lambda^{c,r^*\perp}_{\text{NDE,obs}}$ with the orthocomplement of any model for which the NDE assumption only holds for a strict subset of $\{\bar{C}_K, \bar{R}^*_K, Y^d\}$, since $V_{\text{nde-ipw},g} - \Pi[V_{\text{nde-ipw},g}|\Lambda_{\text{ancillary}}]$ explicitly depends on each of $Y^d, \bar{C}_K$, and $\bar{R}^*_K$. Indeed, unless the NDE$(\bar{C}_K, \bar{R}^*_K)$ assumption holds, $V_{\text{nde-ipw},g} - \Pi[V_{\text{nde-ipw},g}|\Lambda_{\text{ancillary}}]$ is not a function of the observed data. (So neither is $\widehat{V}_{\text{nde-ipw},g} - \widehat{\Pi}[\widehat{V}_{\text{nde-ipw},g}|\Lambda_{\text{ancillary}}]$, where $\widehat{\Pi}$ denotes the projection onto the space $\Lambda_{\text{ancillary}}$ with all nuisance functions replaced by their estimates.)
2. Next suppose, we are only willing to assume model $\mathbb{M}_{\text{NDE}}(Y^d)$ that just imposes the NDE(Y^d) and modified ID assumptions. Our goal is again to estimate θ_g for a regime g_t that depends on $\bar{r}_t$, $\bar{c}_t$, and $\bar{s}_{t-1}$ so $\widehat{\theta}_{\text{nde-ipw},g}$ (and hence $\widetilde{\theta}_{\text{nde-ipw},g}$) is not a statistic. Nevertheless we can still improve upon the efficiency of $\widetilde{\theta}_{\text{ipw},g} = \mathbb{P}_n[\widehat{V}_{\text{ipw},g}]$. We first give the oracle version of the procedure. Let $U_{\text{ipw},g}(\theta) = V_{\text{ipw},g} - \Pi[V_{\text{ipw},g}|\Lambda_{\text{ancillary}}] - \theta$ so $U_{\text{ipw},g}(\theta_g) = \text{IF}_{\text{ipw},g}$ is an influence function for θ_g in $\mathbb{M}_{\text{NDE}}(Y^d)$. It follows that $\widehat{\theta}_{\text{ipw},g}(b_{\text{opt,ipw},g})$ solving $\mathbb{P}_n[U_{\text{ipw}}(\theta, b_{\text{opt,ipw},g})] \equiv$

$\mathbb{P}_n[U_{\text{ipw},g}(\theta) - \Pi[U_{\text{ipw},g}(\theta)|\Lambda^{\perp}_{\text{NDE}}]] = 0$ is semiparametric efficient in model in $\mathbb{M}_{\text{NDE}}(Y^d)$, as its influence function is EIF $=$ $\text{IF}_{\text{ipw},g} - \Pi[\text{IF}_{\text{ipw},g}|\Lambda^{\perp}_{\text{NDE}}]$ under the model. Note that

$$U_{\text{ipw},g}(\theta) - \Pi[U_{\text{ipw},g}(\theta)|\Lambda^{\perp}_{\text{NDE}}] = V_{\text{ipw},g} - \theta - \Pi[V_{\text{ipw},g}|\Lambda_{\text{ancillary}}] - \Pi[V_{\text{ipw},g}|\Lambda^{\perp}_{\text{NDE}}],$$

as neither $\Pi[V_{\text{ipw},g}|\Lambda_{\text{ancillary}}]$ nor θ has a projection on $\Lambda^{\perp}_{\text{NDE}}$. Hence $\widehat{\theta}_{\text{ipw},g}(b_{\text{opt,ipw},g}) = \mathbb{P}_n\{V_{\text{ipw},g} - \Pi[V_{\text{ipw},g}|\Lambda_{\text{ancillary}}] - \Pi[V_{\text{ipw},g}|\Lambda^{\perp}_{\text{NDE}}]\}$. We can obtain a feasible version $\widetilde{\theta}_{\text{ipw},g}(\widehat{b}_{\text{opt,ipw},g})$ of $\widehat{\theta}_{\text{ipw},g}(b_{\text{opt,ipw},g})$ using the analogue of the algorithm proposed in Section 6.

3. Suppose we again assume model $\mathbb{M}_{\text{NDE,obs}}(\bar{C}_K, \bar{R}^*_K, Y^d)$ is true so both $\widetilde{\theta}_{\text{ipw},g} = \mathbb{P}_n[\widehat{V}_{\text{ipw},g}]$ and $\widetilde{\theta}_{\text{nde-ipw},g} = \mathbb{P}_n[\widehat{V}_{\text{nde-ipw},g}]$ are consistent for θ_g. Furthermore $\widetilde{\theta}_{\text{ipw},g} - \widehat{\Pi}[V_{\text{ipw},g}|\Lambda_{\text{ancillary}}]$ and $\widetilde{\theta}_{\text{nde-ipw},g} - \widehat{\Pi}[V_{\text{nde-ipw},g}|\Lambda_{\text{ancillary}}]$ are RAL with the influence functions $\text{IF}_{\text{ipw},g}$ and $\text{IF}^{c,r^*}_{\text{nde-ipw},g}$ for θ_g. Recall that in Caniglia et al. (2019), $\widetilde{\theta}_{\text{nde-ipw},g}$ was 50 times as efficient as $\widetilde{\theta}_{\text{ipw},g}$. Now the asymptotic variances of the single sample estimators $\widetilde{\theta}_{\text{ipw},g}$ and $\widetilde{\theta}_{\text{nde-ipw},g}$ used by Caniglia et al. (2019) approximate $\text{var}\left(\text{IF}_{\text{ipw},g}\right)$ and $\text{var}\left(\text{IF}^{c,r^*}_{\text{nde-ipw},g}\right)$, because flexible, high-dimensional models were used in estimating Π_t and W_0. But (i) neither $\text{var}\left(\text{IF}_{\text{ipw},g}\right)$ nor $\text{var}\left(\text{IF}^{c,r^*}_{\text{nde-ipw},g}\right)$ dominates the other at all laws in the model and both exceed the semiparametric variance bound of model $\mathbb{M}_{\text{NDE,obs}}(\bar{C}_K, \bar{R}^*_K, Y^d)$. Rather the oracle estimators $\widehat{\theta}_{\text{ipw},g}(b^{c,r^*}_{\text{opt,ipw},g}) = \mathbb{P}_n\{V_{\text{ipw},g} - \Pi[V_{\text{ipw},g}|\Lambda_{\text{ancillary}}] - \Pi[V_{\text{ipw},g}|\Lambda^{c,r^*\perp}_{\text{NDE,obs}}]$ solving $\mathbb{P}_n[U_{\text{ipw},g}(\theta) - \Pi[U_{\text{ipw},g}(\theta)|\Lambda^{c,r^*\perp}_{\text{NDE,obs}}]] = 0$ and $\widehat{\theta}_{\text{nde-ipw},g}(b^{c,r^*}_{\text{opt,nde-ipw},g}) = \mathbb{P}_n\{V_{\text{nde-ipw},g} - \Pi[V_{\text{nde-ipw},g}|\Lambda_{\text{ancillary}}] - \Pi[V_{\text{nde-ipw},g}|\Lambda^{c,r^*\perp}_{\text{NDE,obs}}]\}$ solving $\mathbb{P}_n[U_{\text{nde-ipw},g}(\theta) - \Pi[U_{\text{nde-ipw},g}(\theta)|\Lambda^{c,r^*\perp}_{\text{NDE,obs}}]] = 0$ are semiparametric efficient in $\mathbb{M}_{\text{NDE,obs}}(\bar{C}_K, \bar{R}^*_K, Y^d)$, with common influence function $\text{EIF}^{c,r^*}_{\text{obs}} = \text{IF}_{\text{ipw},g} - \Pi[\text{IF}_{\text{ipw},g}|\Lambda^{c,r^*\perp}_{\text{NDE,obs}}] = \text{IF}^{c,r^*}_{\text{nde-ipw},g} - \Pi[\text{IF}^{c,r^*}_{\text{nde-ipw},g}|\Lambda^{c,r^*\perp}_{\text{NDE,obs}}]$. Again the algorithm in Section 6 can be used to obtain feasible efficient estimators $\widetilde{\theta}_{\text{ipw},g}(\widehat{b}^{c,r^*}_{\text{opt,ipw},g})$ and $\widetilde{\theta}_{\text{nde-ipw},g}(\widehat{b}^{c,r^*}_{\text{opt,nde-ipw},g})$ with influence function $\text{EIF}^{c,r^*}_{\text{obs}}$.

9. Simulation Studies

9.1. A Simple DGP Related to Caniglia et al. (2019)

This simulation study explores in greater depth DGPs under which $\widetilde{\theta}_{\text{nde-ipw},g}$ is much more efficient than $\widetilde{\theta}_{\text{ipw},g}$. The study also relates these estimators to estimators of the optimal value function $\text{E}[Y_{g^{\text{opt}}}]$ based on opt-SNMMs considered in Section 3.

We consider a much simplified version of the HIV study of Caniglia et al. (2019) in which $K = 1$ with time-ordered random variables: $A_0, R^*_1, R_1, S_1, Y^d, Y$. Here A_0 is the indicator of having an HIV RNA test; R_1 is the test result: R_1 is ? if $A_0 = 0$ and is R^*_1 if $A_0 = 1$; R^*_1 is the CD4 count dichotomized as $R^*_1 = 0$ if low and $R^*_1 = 1$ otherwise; S_1 is the indicator of treatment (i.e., of switching to second line ART); Y^d is a health utility that is a decreasing function of HIV RNA levels at time $K+1$; $Y := Y^d - c^*A_0$ with $c^* = 3$ is the overall observed utility incorporating the cost of testing. Since we are including the cost of testing we replace Y by $Y_g = Y + \mathbb{1}[A_0 = 1, A_{0,g} = 0]$ in defining $\widetilde{\theta}_{\text{ipw},g}$ and $\widetilde{\theta}_{\text{nde-ipw},g}$ as discussed in Remark 5. Y_g is just the health utility Y^d for any regime with $A_{0,g} \equiv 0$.

Our DGP was designed such that (i) the NDE(R^*_1, Y^d) and modified ID assumptions hold and (ii) $g^{\text{opt}} = (a^{\text{opt}}_0 = 0, s^{\text{opt}}_1 = 1)$ and hence $\theta_{g^{\text{opt}}} = \text{E}[Y_{a_0=0,s_1=1}] \equiv \text{E}[Y^d_{a_0=0,s_1=1}]$. Simulations with more complex DGPs with g^{opt} depending on R can be found in Appendix A.9.

The DGP is

- $A_0 \sim$ Bernoulli(ρ), with $\rho \in \{0.1, 0.2, \ldots, 0.8, 0.9, 0.95, 0.99\}$;
- $R^*_1 \sim$ Bernoulli(0.5);
- $R_1 = R^*_1$ if $A_0 = 1$ and ? otherwise;
- $S_1 \sim$ Bernoulli$(0.15)\mathbb{1}\{A_0 = 1, R_1 = 0\}$ + Bernoulli$(0.85)\mathbb{1}\{A_0 = 1, R_1 = 1\}$+ Bernoulli$(0.15)\mathbb{1}\{A_0 = 0, R_1 =?\}$;
- $Y^d \sim N(-5R^*_1 + 2S_1R^*_1 - 0.1S_1(1 - R^*_1), 1)$;
- $Y := Y^d - c^*A_0$ with $c^* = 3$.

Since A_0 is neither a direct cause of Y nor of R^*_1, the NDE (R^*_1, Y^d) assumption holds. The sample size was $n = 2.5 \times 10^4$ or 5×10^4. We computed the "truth" based on a single simulated dataset with sample size 10^7. *Monte Carlo* summary statistics are based on 100 replications. The R codes for the simulation studies are provided in the online supplementary materials.

We note several other properties of our DGP. Whenever $K = 1$, it follows from the explicit representation of $\Lambda^{r^*\perp}_{\text{NDE,obs}}$ given in Appendix A.7 that the ortho-complements to the observed data tangent spaces in models $\mathbb{M}_{\text{NDE}}(Y^d)$ and $\mathbb{M}_{\text{NDE,obs}}(\bar{R}^*_K, Y^d)$ are equal; that is, $\Lambda^{\perp}_{\text{NDE}} = \Lambda^{r^*\perp}_{\text{NDE,obs}}$ because the latter cannot be a function of R_1 or R^*_1. We further note that, as $\rho \to 1$, the probability of the event $\{A_0 = 1, A_{0,g^{\text{opt}}} = 0\}$ that testing occurred too early in the observed data increases to 1, mimicking what happened in Caniglia et al. (2019) discussed in Section 8. We are thus interested in precisely how ρ affects the efficiency to be gained by imposing the NDE assumption.

We did not need or use sample splitting in our analyses because A_0, R_1, S_1 are discrete with few levels; we estimated their joint distribution by their empirical distribution. We approximate $\Lambda^{\perp}_{\text{NDE}}$ by Ω using natural spline transformations $\boldsymbol{\varphi}$ with $\xi = 6$. More precisely, in our simulation, $\Omega = \{d_0 T_{b^*,0} : d_0 \in \mathbb{R}^{2\xi}\}$, with $b^*_0(\bar{H}_0, \underline{S}_0, Y^d) \equiv b^*_0(S_1, Y^d) \equiv (S_1\boldsymbol{\varphi}(Y^d)^\top, (1 - S_1)\boldsymbol{\varphi}(Y^d)^\top)^\top$. Hence, for any U, $\Pi[U|\Omega] = d^{*}_0 T_{b^*,0} \equiv T_{b_{\text{sub}},0}$ (following the notation in Corollary 1), with $d^*_0 = \text{E}[UT^\top_{b^*,0}]\{\text{E}[T_{b^*,0}T^\top_{b^*,0}]\}^{-1}$ and $T_{b^*,0}$ as in Equation (10), with b_t replaced by $b^*_0(S_1, Y^d)$. The feasible version of $\Pi[U|\Omega]$ in our simulation is $\widehat{T}_{b_{\text{sub}},0} := \widehat{\Pi}[U|\Omega] = \widetilde{d}^*_0\widehat{T}_{b^*,0}$, where $\widetilde{d}^*_0 = \mathbb{P}_n[UT^\top_{b^*,0}]\{\mathbb{P}_n[\widehat{T}_{b^*,0}\widehat{T}^\top_{b^*,0}]\}^{-1}$ and $\widehat{T}_{b^*,0}$ is $T_{b^*,0}$ with all unknown nuisance functions replaced by their empirical version as described above.

We then used the entire sample to estimate $\theta_{g^{\text{opt}}}$ with each of the following estimators in our simulation:

$$\widetilde{\theta}_{\text{ipw},g^{\text{opt}}} := \mathbb{P}_n[\widehat{V}_{\text{ipw},g^{\text{opt}}}], \widehat{V}_{\text{ipw},g^{\text{opt}}} = \frac{(1-A_0)S_1 Y}{\widehat{\Pr}(A_0=0)\widehat{\Pr}(S_1=1|A_0=0,R_1)},$$
$$\widetilde{\theta}_{\text{nde-ipw},g^{\text{opt}}} := \mathbb{P}_n[\widehat{V}_{\text{nde-ipw},g^{\text{opt}}}], \widehat{V}_{\text{nde-ipw},g^{\text{opt}}} = \frac{S_1 Y_{g^{\text{opt}}}}{\widehat{\Pr}(S_1=1|A_0,R_1)},$$
$$\widetilde{\theta}_{\text{ipw},g^{\text{opt}}}(\widehat{b}_{\text{sub,ipw},g^{\text{opt}}}) := \widetilde{\theta}_{\text{ipw},g^{\text{opt}}} - \mathbb{P}_n\{\widehat{\Pi}[\widehat{V}_{\text{ipw},g^{\text{opt}}} - \widetilde{\theta}_{\text{ipw},g^{\text{opt}}}|\Omega]\},$$
$$\widetilde{\theta}_{\text{nde-ipw},g^{\text{opt}}}(\widehat{b}_{\text{sub,nde-ipw},g^{\text{opt}}}) := \widetilde{\theta}_{\text{nde-ipw},g^{\text{opt}}} - \mathbb{P}_n\{\widehat{\Pi}[\widehat{V}_{\text{nde-ipw},g^{\text{opt}}} - \widetilde{\theta}_{\text{nde-ipw},g^{\text{opt}}}|\Omega]\},$$
$$\widetilde{\theta}_{\text{SNMM},g^{\text{opt}}}(\widetilde{\Psi}) := \mathbb{P}_n[\Delta_0(\widetilde{\Psi}_0;\underline{\widetilde{\Psi}}_1) + \gamma_0(A_0^{\text{opt}}(\underline{\widetilde{\Psi}}_0);\widetilde{\Psi}_0)],$$
$$\widetilde{\theta}_{\text{SNMM},g^{\text{opt}}}(\widetilde{\Psi};\widehat{b}_{\text{sub},g^{\text{opt}}}) := \widetilde{\theta}_{\text{SNMM},g^{\text{opt}}}(\widetilde{\Psi}) - \mathbb{P}_n\{\widehat{\Pi}[\text{IF}_{\text{E}[\widehat{U}_{\text{SNMM},g^{\text{opt}}}]}(\widetilde{\theta}_{\text{SNMM},g^{\text{opt}}}(\widetilde{\Psi}),\widetilde{\Psi})|\Omega]\},$$

where the oracle version of $\widehat{U}_{\text{SNMM},g^{\text{opt}}}(\widetilde{\theta}_{\text{SNMM},g^{\text{opt}}}(\widetilde{\Psi}),\widetilde{\Psi}) \equiv \Delta_0(\widetilde{\Psi}_0;\underline{\widetilde{\Psi}}_1) + \gamma_0(A_0^{\text{opt}}(\underline{\widetilde{\Psi}}_0);\widetilde{\Psi}_0) - \widetilde{\theta}_{\text{SNMM},g^{\text{opt}}}(\widetilde{\Psi})$ is defined in Equation (7). For the last two estimators, we fit a saturated opt-SNMM. In a saturated opt-SNMM, Ψ^* is 7 dimensional with $6 = 2 \times 3$ parameters for the effect of S_1 on Y_g within levels of $\bar{H}_1 = (A_0, R_1)$ and 1 parameter for the effect of A_0.

9.2. Simulation Results

In the simulation, at both sample sizes $n = 2.5 \times 10^4$ or 5×10^4, the optimal regime estimate from our saturated opt-SNMM model was equal to the true optimal regime $g^{\text{opt}} = (a_0 = 0, s_1 = 1)$ in each of the 100 replications. Because $\widetilde{\theta}_{\text{SNMM},g^{\text{opt}}}(\widetilde{\Psi})$ and $\widetilde{\theta}_{\text{ipw},g^{\text{opt}}}$ are both Fisher consistent nonparametric MLEs of $\theta_{g^{\text{opt}}}$ in the absence of the NDE assumption, we expected and empirically verified that they are algebraically identical in each of the 100 replications. Furthermore since $\widetilde{\theta}_{\text{SNMM},g^{\text{opt}}}(\widetilde{\Psi};\widehat{b}_{\text{sub},g^{\text{opt}}})$ and $\widetilde{\theta}_{\text{ipw},g^{\text{opt}}}(\widehat{b}_{\text{sub,ipw},g^{\text{opt}}})$ are then computed by subtracting the empirical projections on the same space Ω, we expected and empirically verified that they too are algebraically identical. As a result we do not separately report results for $\widetilde{\theta}_{\text{SNMM},g^{\text{opt}}}(\widetilde{\Psi})$ or $\widetilde{\theta}_{\text{SNMM},g^{\text{opt}}}(\widetilde{\Psi};\widehat{b}_{\text{sub},g^{\text{opt}}})$.

The results of the simulation study are in Figure 1. The results are consistent with the intuition we gained in Section 8. In Figure 1, we plot the following as a function of ρ: (1) in the left panel, the Monte Carlo variances of $\widetilde{\theta}_{\text{ipw},g^{\text{opt}}}$ (black), $\widetilde{\theta}_{\text{nde-ipw},g^{\text{opt}}}$ (blue), $\widetilde{\theta}_{\text{ipw},g^{\text{opt}}}(\widehat{b}_{\text{sub,ipw},g^{\text{opt}}})$ (red) and $\widetilde{\theta}_{\text{nde-ipw},g^{\text{opt}}}(\widehat{b}_{\text{sub, nde-ipw},g^{\text{opt}}})$ (green); and (2) in the right panel, the inverse relative efficiencies (REs) of $\widetilde{\theta}_{\text{nde-ipw},g^{\text{opt}}}$ (blue), $\widetilde{\theta}_{\text{ipw},g^{\text{opt}}}(\widehat{b}_{\text{sub,ipw},g^{\text{opt}}})$ (red) and $\widetilde{\theta}_{\text{nde-ipw},g^{\text{opt}}}(\widehat{b}_{\text{sub, nde-ipw},g^{\text{opt}}})$ (green) against $\widetilde{\theta}_{\text{ipw},g^{\text{opt}}}$.

We now summarize the results of Figure 1:

- In the left panel of Figure 1, the Monte Carlo variance of $\widetilde{\theta}_{\text{ipw},g^{\text{opt}}}$ increases dramatically as ρ (and thus the probability of the event $[A_0 = 1, A_{0,g^{\text{opt}}} = 0]$) approaches 1. The variance increases because all subjects i with $A_{0,i} = 1$ are censored so $\widehat{V}_{\text{ipw},g^{\text{opt}},i} = 0$. In contrast, the Monte Carlo variances of the three estimators that leverage the NDE assumption are stable as ρ varies and cannot be easily distinguished from one another at this scale.
- In the right panel of Figure 1, we compare the inverse RE of the three estimators leveraging the NDE assumption relative to the estimator $\widetilde{\theta}_{\text{ipw},g^{\text{opt}}}$. As ρ approaches 1, the inverse relative efficiency of the three estimators decreases to 0.

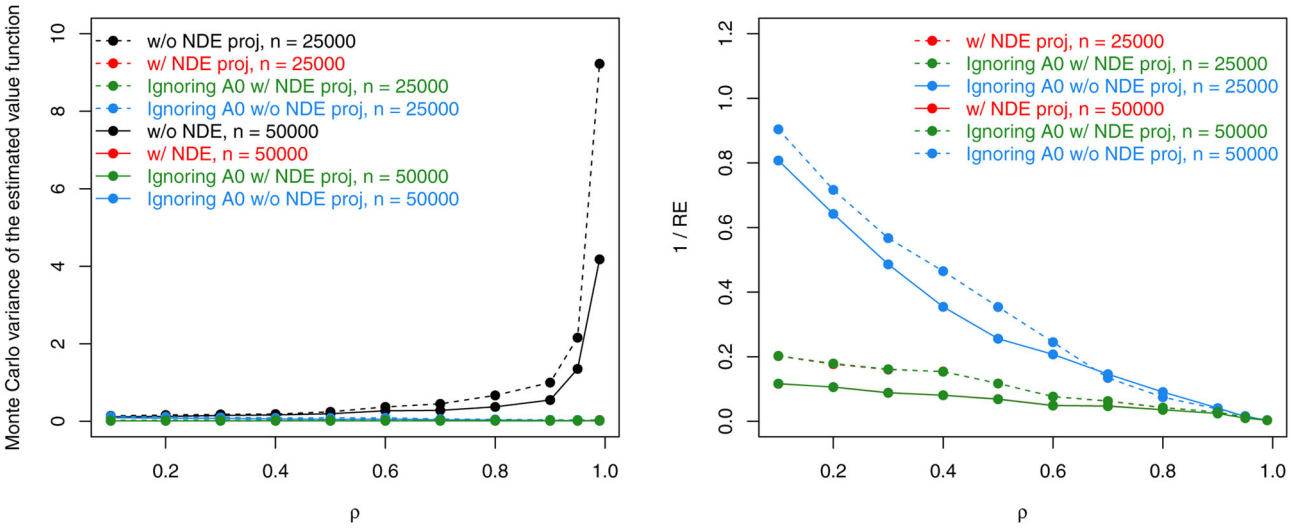

Figure 1. The Monte Carlo variances (left panel) of $\widetilde{\theta}_{\text{ipw},g^{\text{opt}}}$ (black), $\widetilde{\theta}_{\text{nde-ipw},g^{\text{opt}}}$ (blue), $\widetilde{\theta}_{\text{ipw},g^{\text{opt}}}(\widehat{b}_{\text{sub,ipw},g^{\text{opt}}})$ (red), and $\widetilde{\theta}_{\text{nde-ipw},g^{\text{opt}}}(\widehat{b}_{\text{sub,nde-ipw},g^{\text{opt}}})$ (green), and inverse REs (right panel) of $\widetilde{\theta}_{\text{nde-ipw},g^{\text{opt}}}$ (blue), $\widetilde{\theta}_{\text{ipw},g^{\text{opt}}}(\widehat{b}_{\text{sub,ipw},g^{\text{opt}}})$ (red), and $\widetilde{\theta}_{\text{nde-ipw},g^{\text{opt}}}(\widehat{b}_{\text{sub,nde-ipw},g^{\text{opt}}})$ (green) against $\widetilde{\theta}_{\text{ipw},g^{\text{opt}}}$ over different ρ's. The red and green lines lie on top of one another; due to scaling, the blue and green lines are indistinguishable in the left panel.

When ρ is far from 1, the estimators $\widetilde{\theta}_{\text{ipw},g^{\text{opt}}}(\widehat{b}_{\text{sub,ipw},g^{\text{opt}}})$ and $\widetilde{\theta}_{\text{nde-ipw},g^{\text{opt}}}(\widehat{b}_{\text{sub,nde-ipw},g^{\text{opt}}})$ (red and green curves, which are nearly on top of each other) are more efficient than $\widetilde{\theta}_{\text{nde-ipw},g^{\text{opt}}}$ because the former two are nearly semiparametric efficient in the model $\mathbb{M}_{\text{NDE}}(Y^d)$, as they are residuals from the projection of an influence function onto a large subspace Ω of the ortho-complement $\Lambda^{\perp}_{\text{NDE}}$ to the tangent space of $\mathbb{M}_{\text{NDE}}(Y^d)$. However, as ρ approaches 1, the variance of $\widetilde{\theta}_{\text{nde-ipw},g^{\text{opt}}}$ converges from above to that of $\widetilde{\theta}_{\text{ipw},g^{\text{opt}}}(\widehat{b}_{\text{sub,ipw},g^{\text{opt}}})$ and $\widetilde{\theta}_{\text{nde-ipw},g^{\text{opt}}}(\widehat{b}_{\text{sub,nde-ipw},g^{\text{opt}}})$, as the ortho-complement to the tangent space of model $\mathbb{M}_{\text{NDE}}(Y^d)$ becomes degenerate at $\rho = 1$. In fact, when $\rho = 1$ (not shown in the plot), we can show the following: $\widetilde{\theta}_{\text{ipw},g^{\text{opt}}}$ and, thus $\widetilde{\theta}_{\text{ipw},g^{\text{opt}}}(\widehat{b}_{\text{sub,ipw},g^{\text{opt}}})$ are undefined and thus inconsistent due to a positivity violation: $A_0 = 0$ with probability 0. In contrast, $\widetilde{\theta}_{\text{nde-ipw},g^{\text{opt}}}$ and $\widetilde{\theta}_{\text{nde-ipw},g^{\text{opt}}}(\widehat{b}_{\text{sub,nde-ipw},g^{\text{opt}}})$ are equal with probability 1 and are RAL. In fact, they are semiparametric efficient for $\theta_{g^{\text{opt}}}$ in model $\mathbb{M}_{\text{NDE}}(Y^d)$ at $\rho = 1$ since $\theta_{g^{\text{opt}}}$ is just identified and so has a unique influence function. Further details of this simulation are presented in Appendix A.9.

In Appendix A.10, we provide a second simulation which shows that adjusting for the NDE assumption in an opt-SNMM can greatly improve the efficiency of estimating the VoI in the context of a cost benefit analysis. As mentioned in Section 1, $\widetilde{\theta}_{\text{nde-ipw},g^{\text{opt}}}$ can be less efficient than the estimator $\widetilde{\theta}_{\text{ipw},g^{\text{opt}}}$ that ignores the NDE assumption. In Appendix A.12, we design a data generating process to illustrate this point. We perform exact calculations of the variance of the oracle estimators $\widehat{\theta}_{\text{ipw},g^{\text{opt}}}$ and $\widehat{\theta}_{\text{nde-ipw},g^{\text{opt}}}$ and show $\text{var}[\widehat{\theta}_{\text{nde-ipw},g^{\text{opt}}}]/\text{var}[\widehat{\theta}_{\text{ipw},g^{\text{opt}}}] > 1$. We actually show this ratio can be made arbitrarily large by varying a parameter of the DGP.

10. Discussion

To conclude, we discuss five open problems. The optimal choice of $\boldsymbol{q}^{\text{opt}}$ for the user-supplied function $\boldsymbol{q}$ is the unique vector function $\boldsymbol{q}^{\text{opt}}$ satisfying for all $\boldsymbol{q} = \{q_t(\bar{H}_t, s_t, a_t); t = 0, \ldots, K\}$,

$$\text{E}\left[\frac{\partial \text{E}\left[\mathbb{U}(\boldsymbol{q}, \Psi^*)\right]}{\partial \Psi^\top}\right] = \text{E}[\mathbb{U}(\boldsymbol{q}, \Psi^*)\mathbb{U}(\boldsymbol{q}^{\text{opt}}, b_{\text{opt}}(\boldsymbol{q}^{\text{opt}}), \Psi^*)^\top].$$

Furthermore $\left\{\text{E}\left[\mathbb{U}(\boldsymbol{q}^{\text{opt}}, b_{\text{opt}}(\boldsymbol{q}^{\text{opt}}), \Psi^*)\mathbb{U}(\boldsymbol{q}^{\text{opt}}, b_{\text{opt}}(\boldsymbol{q}^{\text{opt}}), \Psi^*)^\top\right]\right\}^{-1}$ is the semiparametric variance bound for Ψ^* in model $\mathbb{M}$. We did not explore estimation of $\boldsymbol{q}^{\text{opt}}$ because, as discussed earlier, in practice, users of opt-SNMM have chosen to employ heuristic choices of $\boldsymbol{q}$ in their analyses. Nonetheless, it would be interesting to study the estimation with $\boldsymbol{q}^{\text{opt}}$ in future research. A second open problem is to develop multiply robust estimators of the parameters of an opt-SNMM under the NDE assumption to provide even more robustness than that obtained by the doubly robust estimators proposed in the current article (Luedtke et al. 2017; Rotnitzky, Robins, and Babino 2017).

A third open problem is to extend our methodology to the case where either (i) the dimension of the parameter Ψ^* of the opt-SNMM is allowed to grow with the sample size or (ii) the optimal blip functions $\gamma_t^{g^{\text{opt}}}(\bar{H}_t, s_t, a_t)$ are modeled nonparametrically under (e.g., smoothness) restrictions on their complexity. This problem is important because our estimator of the optimal regime is not robust to misspecification of the opt-SNMM model. A fourth important open problem is the extension of our methods to data generated under exceptional laws as defined in Robins (2004).

We conclude by discussing a fifth open problem. Consider the setting discussed earlier under which positivity fails because $\Pr(A_t = 1) = 1$ for all t but other aspects of positivity, consistency and sequential exchangeability continue to hold. In this setting, the parameter vector Ψ^* of a saturated opt-SNMM is not identified if the NDE assumption fails to hold. In fact, under the NDE $(\bar{L}_K, \bar{R}_K, Y^d)$ assumption, all g for which θ_g was identified under positivity remain identified; similarly, since Ψ^* is identified, g^{opt} and $\theta_{g^{\text{opt}}}$ remain identified under this NDE assumption. However, by lack of identification in the absence of the NDE assumption, the methodology of this article cannot be used to estimate Ψ^*. It remains open how to efficiently estimate Ψ^* and thus g^{opt} and $\theta_{g^{\text{opt}}}$ in this setting under the NDE assumption. Preliminary investigations indicate that estimation of g^{opt} by dynamic programming is not possible. If true, then any algorithms that compute or estimate g^{opt} may well be computationally intractable.

Supplementary materials

Appendix and R codes are included in the online supplementary materials.

Acknowledgments

We thank Michael R. Kosorok and two anonymous referees for helpful comments. This work was motivated by a talk given by James M. Robins at ENAR 2015 entitled "How to increase efficiency of estimation when a test used to decide treatment has no direct effect on the outcome" in a session organized by Michael R. Kosorok.

Funding

Lin Liu and James M. Robins were supported by the U.S. Office of Naval Research grant N000141912446, and National Institutes of Health (NIH) awards R01 AG057869 and R01 AI127271. Lin Liu was also partially sponsored by Shanghai Pujiang Program Research grant 20PJ1408900 and Shanghai Jiao Tong University Start-up Grant WF220441912. The computations were run on the FASRC Cannon cluster supported by the FAS Division of Science Research Computing Group at Harvard University.

References

Bang, H., and Robins, J. M. (2005), "Doubly Robust Estimation in Missing Data and Causal Inference Models," *Biometrics*, 61, 962–973. [3]

Bellman, R. (1952), "On the Theory of Dynamic Programming," *Proceedings of the National Academy of Sciences of the United States of America*, 38, 716. [5]

Caniglia, E. C., Robins, J. M., Cain, L. E., Sabin, C., Logan, R., Abgrall, S., Mugavero, M. J., Hernández-Díaz, S., Meyer, L., Seng, R., and Drozd, D. R. (2019), "Emulating a Trial of Joint Dynamic Strategies: An Application to Monitoring and Treatment of HIV-Positive Individuals," *Statistics in Medicine*, 38, 2428–2446. [1,3,4,10,11,13]

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., and Robins, J. (2018), "Double/Debiased Machine Learning for Treatment and Structural Parameters," *The Econometrics Journal*, 21, C1–C68. [9]

DART Trial Team (2010), "Routine Versus Clinically Driven Laboratory Monitoring of HIV Antiretroviral Therapy in Africa (DART): A Randomised Non-Inferiority Trial," *The Lancet*, 375, 123–131. [1]

Ford, D., Robins, J. M., Petersen, M. L., Gibb, D. M., Gilks, C. F., Mugyenyi, P., Grosskurth, H., Hakim, J., Katabira, E., Babiker, A. G., and Walker, A. S. (2015), "The Impact of Different CD4 Cell-Count Monitoring and Switching Strategies on Mortality in HIV-Infected African Adults on Antiretroviral Therapy: An Application of Dynamic Marginal Structural Models," *American Journal of Epidemiology*, 182, 633–643. [2]

Gould, J. P. (1974), "Risk, Stochastic Preference, and the Value of Information," *Journal of Economic Theory*, 8, 64–84. [2]

Hernán, M. A., Robins, J. M., and García Rodríguez, L. A. (2005), "Discussion on 'Statistical issues arising in the Women's Health Initiative' by Prentice RL et al.," *Biometrics*, 61, 922–930. [3]

Hilton, R. W. (1981), "The Determinants of Information Value: Synthesizing Some General Results," *Management Science*, 27, 57–64. [2]

Kallus, N., Mao, X., and Uehara, M. (2019), "Localized Debiased Machine Learning: Efficient Estimation of Quantile Treatment Effects, Conditional Value at Risk, and Beyond," arXiv no. 1912.12945. [9]

Krahn, M. D., Mahoney, J. E., Eckman, M. H., Trachtenberg, J., Pauker, S. G., and Detsky, A. S. (1994), "Screening for Prostate Cancer: A Decision Analytic View," *JAMA*, 272, 773–780. [2]

Kreif, N., Sofrygin, O., Schmittdiel, J. A., Adams, A. S., Grant, R. W., Zhu, Z., van der Laan, M. J., and Neugebauer, R. (2020), "Exploiting Nonsystematic Covariate Monitoring to Broaden the Scope of Evidence About the Causal Effects of Adaptive Treatment Strategies," *Biometrics*. [3]

Lara, A. M., Kigozi, J., Amurwon, J., Muchabaiwa, L., Wakaholi, B. N., Mota, R. E. M., Walker, A. S., Kasirye, R., Ssali, F., Reid, A., and Grosskurth, H. (2012), "Cost Effectiveness Analysis of Clinically Driven Versus Routine Laboratory Monitoring of Antiretroviral Therapy in Uganda and Zimbabwe," *PLoS ONE*, 7, e33672. [2]

LaValle, I. H. (1968a), "On Cash Equivalents and Information Evaluation in Decisions Under Uncertainty Part I: Basic Theory," *Journal of the American Statistical Association*, 63, 252–276. [2]

——— (1968b), "On Cash Equivalents and Information Evaluation in Decisions Under Uncertainty Part II: Incremental Information Decisions," *Journal of the American Statistical Association*, 63, 277–284. [2]

Luedtke, A. R., Sofrygin, O., van der Laan, M. J., and Carone, M. (2017), "Sequential Double Robustness in Right-Censored Longitudinal Models," arXiv no. 1705.02459. [15]

Murphy, S. A. (2003), "Optimal Dynamic Treatment Regimes," *Journal of the Royal Statistical Society*, Series B, 65, 331–355. [2]

Mushlin, A. I., and Fintor, L. (1992), "Is Screening for Breast Cancer Cost-Effective?," *Cancer*, 69, 1957–1962. [2]

Neugebauer, R., Schmittdiel, J. A., Adams, A. S., Grant, R. W., and van der Laan, M. J. (2017), "Identification of the Joint Effect of a Dynamic Treatment Intervention and a Stochastic Monitoring Intervention Under the No Direct Effect Assumption," *Journal of Causal Inference*, 5, 1–44. [3]

Orellana, L., Rotnitzky, A., and Robins, J. M. (2010a), "Dynamic Regime Marginal Structural Mean Models for Estimation of Optimal Dynamic Treatment Regimes, Part I: Main Content," *The International Journal of Biostatistics*, 6, 1–48. [3]

——— (2010b), "Dynamic Regime Marginal Structural Mean Models for Estimation of Optimal Dynamic Treatment Regimes, Part II: Proofs of Results," *The International Journal of Biostatistics*, 6, 1–17. [3]

Robins, J. M. (1986), "A New Approach to Causal Inference in Mortality Studies With a Sustained Exposure Period—Application to Control of the Healthy Worker Survivor Effect," *Mathematical Modelling*, 7, 1393–1512. [4,5]

——— (1987), "Addendum to 'A New Approach to Causal Inference in Mortality Studies With a Sustained Exposure Period—Application to Control of the Healthy Worker Survivor Effect'," *Computers & Mathematics With Applications*, 14, 923–945. [4]

——— (1999), "Testing and Estimation of Direct Effects by Reparameterizing Directed Acyclic Graphs With Structural Nested Models," in *Computation, Causation, and Discovery*, eds. C. Glymour and G. Cooper, Menlo Park, CA: AAAI Press, pp. 349–405. [5,7,9,10]

——— (2000), "Marginal Structural Models Versus Structural Nested Models as Tools for Causal Inference," in *Statistical Models in Epidemiology, the Environment, and Clinical Trials*, eds. M. E. Halloran and D. Berry, New York: Springer, pp. 95–133. [2]

——— (2004), "Optimal Structural Nested Models for Optimal Sequential Decisions," in *Proceedings of the Second Seattle Symposium in Biostatistics*, Springer, pp. 189–326. [2,4,5,6,15]

Robins, J. M., Orellana, L., and Rotnitzky, A. (2008), "Estimation and Extrapolation of Optimal Treatment and Testing Strategies," *Statistics in Medicine*, 27, 4678–4721. [3,10,11,12]

Robins, J. M., and Rotnitzky, A. (1995), "Semiparametric Efficiency in Multivariate Regression Models With Missing Data," *Journal of the American Statistical Association*, 90, 122–129.

Rotnitzky, A., Robins, J. M., and Babino, L. (2017), "On the Multiply Robust Estimation of the Mean of the g-Functional," arXiv no. 1705.08582. [15]

Smucler, E., Rotnitzky, A., and Robins, J. M. (2019), "A Unifying Approach for Doubly-Robust ℓ_1 Regularized Estimation of Causal Contrasts," arXiv no. 1904.03737. [9]

van der Laan, M. J., and Petersen, M. L. (2007), "Causal Effect Models for Realistic Individualized Treatment and Intention to Treat Rules," *The International Journal of Biostatistics*, 3, 1–55. [3]

van der Vaart, A. W. (1998), *Asymptotic Statistics* (Vol. 3), Cambridge: Cambridge University Press. [6]

van der Vaart, A. W., and Wellner, J. (1996), *Weak Convergence and Empirical Processes: With Applications to Statistics*, New York: Springer. [9]

Vansteelandt, S., and Joffe, M. (2014), "Structural Nested Models and G-Estimation: The Partially Realized Promise," *Statistical Science*, 29, 707–731. [2,7]

World Health Organization (2001), "Report of the Commission on Macroeconomics and Health." [2]